

;; TITLE OF INVENTION: METHODS OF ASSAYING FOR G  
;; FILE REFERENCE: MNI-131  
;; CURRENT APPLICATION NUMBER: US/09/880,137  
;; CURRENT FILING DATE: 2001-03-05  
;; PRIOR APPLICATION NUMBER: US 60/186,706  
;; PRIOR FILING DATE: 2000-03-03  
;; NUMBER OF SEQ ID NOS: 8  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 410  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-880-137-3

Query Match 59.0%; Score 36; DB 4; Length 410;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RAQKVQLKLGK 12  
Db 99 RLQERLIKKLGE 110

RESULT 23  
US-09-880-137-1  
; Sequence 1, Application US/09880137  
; Patent No. 6640025  
; GENERAL INFORMATION:  
; APPLICANT: Barstein, Gabriel  
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G  
; FILE REFERENCE: MNI-131  
; CURRENT APPLICATION NUMBER: US/09/880,137  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 60/186,706  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-880-137-1

Query Match 59.0%; Score 36; DB 4; Length 418;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RAQKVQLKLGK 12  
Db 99 RLQERLIKKLGE 110

RESULT 24  
US-09-880-137-2  
; Sequence 2, Application US/09880137  
; Patent No. 6640025  
; GENERAL INFORMATION:  
; APPLICANT: Barstein, Gabriel  
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G  
; FILE REFERENCE: MNI-131  
; CURRENT APPLICATION NUMBER: US/09/880,137  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 60/186,706  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-880-137-2

Query Match 59.0%; Score 36; DB 4; Length 418;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RAQKVQLKLGK 12  
Db 99 RLQERLIKKLGE 110

RESULT 25  
US-09-107-532A-6174  
; Sequence 6174, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6174:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 153 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...153  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6174:

US-09-107-532A-6174

Query Match 57.4%; Score 35; DB 4; Length 153;  
Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AQEKVLKLGK 12  
Db 40 AQEKLVKGGK 50

RESULT 26  
US-09-702-953B-10

; Sequence 10, Application US/09702953B  
; Patent No. 6673897  
; GENERAL INFORMATION:  
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL  
; TITLE OF INVENTION: NOVEL INHIBITORS OF NF-kappaB ACTIVATION  
; FILE REFERENCE: 2676-4554US  
; CURRENT APPLICATION NUMBER: US/09/702,953B  
; CURRENT FILING DATE: 2000-10-31  
; PRIOR APPLICATION NUMBER: PCT/BE99/00055  
; PRIOR FILING DATE: 1999-05-05  
; PRIOR APPLICATION NUMBER: 98201472.2  
; PRIOR FILING DATE: 1998-05-06  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-702-953B-10

Query Match 57.4%; Score 35; DB 4; Length 228;  
Best Local Similarity 63.6%; Pred. No. 1.6e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QEKVLQKLGKA 13  
|||:|:|  
Db 6 QEKELQLANKA 16

RESULT 27  
US-09-880-137-8  
; Sequence 8, Application US/09880137  
; Patent No. 6640025  
; GENERAL INFORMATION:  
; APPLICANT: Berstein, Gabriel  
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G  
; TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS  
; FILE REFERENCE: MNI-131  
; CURRENT APPLICATION NUMBER: US/09/880,137  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 60/186,706  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation  
US-09-880-137-8

Query Match 57.4%; Score 35; DB 4; Length 382;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
|||:|:|:|  
Db 100 RLQDLLEKLGQ 111

RESULT 28  
US-09-107-532A-6398  
; Sequence 6398, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street

; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Denise  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6398:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 399 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (5) LOCATION 1...399  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6398:  
US-09-107-532A-6398

Query Match 57.4%; Score 35; DB 4; Length 399;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
|||:|:|:|  
Db 13 QAREKVKQNGR 24

RESULT 29  
US-09-880-137-4  
; Sequence 4, Application US/09880137  
; Patent No. 6640025  
; GENERAL INFORMATION:  
; APPLICANT: Berstein, Gabriel  
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G  
; TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS  
; FILE REFERENCE: MNI-131  
; CURRENT APPLICATION NUMBER: US/09/880,137  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 60/186,706  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-137-4

Query Match 57.4%; Score 35; DB 4; Length 409;  
Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 Db 100 RLQDRLRLKLGQ 111

## RESULT 30

US-09-880-137-7  
 ; Sequence 7, Application US/09880137  
 ; Patent No. 6640025

## GENERAL INFORMATION:

; APPLICANT: Berstein, Gabriel  
 ; TITLE OF INVENTION: METHODS OF ASSAYING FOR G  
 ; TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS  
 ; FILE REFERENCE: MNI-131  
 ; CURRENT APPLICATION NUMBER: US/09/880,137  
 ; CURRENT FILING DATE: 2001-03-05  
 ; PRIOR APPLICATION NUMBER: US 60/186,706  
 ; PRIOR FILING DATE: 2000-03-03  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 409  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence

## FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation  
 US-09-880-137-7

Query Match 57.4%; Score 35; DB 4; Length 409;  
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 Db 100 RLQDRLRLKLGQ 111

## RESULT 31

US-09-328-352-6321  
 ; Sequence 6321, Application US/09328352  
 ; Patent No. 6562958

## GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 6321  
 ; LENGTH: 468  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii

## US-09-328-352-6321

Query Match 57.4%; Score 35; DB 4; Length 468;  
 Best Local Similarity 53.8%; Pred. No. 3.1e+02;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 Db 364 KAQEEAAQKKEA 376

## RESULT 32

US-09-702-953B-4  
 ; Sequence 4, Application US/09702953B  
 ; Patent No. 6673897

## GENERAL INFORMATION:

; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL  
 ; TITLE OF INVENTION: NOVEL INHIBITORS OF NF-kappaB ACTIVATION  
 ; FILE REFERENCE: 2676-455AUS  
 ; CURRENT APPLICATION NUMBER: US/09/702,953B

; CURRENT FILING DATE: 2000-10-31  
 ; PRIOR APPLICATION NUMBER: PCT/BE99/00055  
 ; PRIOR FILING DATE: 1999-05-05  
 ; PRIOR APPLICATION NUMBER: 98201472.2  
 ; PRIOR FILING DATE: 1998-05-06  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 574  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus

## US-09-702-953B-4

Query Match 57.4%; Score 35; DB 4; Length 574;  
 Best Local Similarity 63.6%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13  
 Db 352 QEKELQRLNKA 362

## RESULT 33

US-09-702-953B-3  
 ; Sequence 3, Application US/09702953B  
 ; Patent No. 6673897

## GENERAL INFORMATION:

; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL  
 ; TITLE OF INVENTION: NOVEL INHIBITORS OF NF-kappaB ACTIVATION  
 ; FILE REFERENCE: 2676-455AUS  
 ; CURRENT APPLICATION NUMBER: US/09/702,953B  
 ; CURRENT FILING DATE: 2000-10-31  
 ; PRIOR APPLICATION NUMBER: PCT/BE99/00055  
 ; PRIOR FILING DATE: 1999-05-05  
 ; PRIOR APPLICATION NUMBER: 98201472.2  
 ; PRIOR FILING DATE: 1998-05-06  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 627  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus

## US-09-702-953B-3

Query Match 57.4%; Score 35; DB 4; Length 627;  
 Best Local Similarity 63.6%; Pred. No. 4.1e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13  
 Db 405 QEKELQRLNKA 415

## RESULT 34

US-09-702-953B-2  
 ; Sequence 2, Application US/09702953B  
 ; Patent No. 6673897

## GENERAL INFORMATION:

; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL  
 ; TITLE OF INVENTION: NOVEL INHIBITORS OF NF-kappaB ACTIVATION  
 ; FILE REFERENCE: 2676-455AUS  
 ; CURRENT APPLICATION NUMBER: US/09/702,953B  
 ; CURRENT FILING DATE: 2000-10-31  
 ; PRIOR APPLICATION NUMBER: PCT/BE99/00055  
 ; PRIOR FILING DATE: 1999-05-05  
 ; PRIOR APPLICATION NUMBER: 98201472.2  
 ; PRIOR FILING DATE: 1998-05-06  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 647  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus

## US-09-702-953B-2

```
US-09-702-953B-2
Query Match      57.4%; Score 35; DB 4; Length 647;
Best Local Similarity 63.6%; Pred. No. 4.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 QEKVLQKLGKA 13
DB      425 QEKVIQLNKA 435

RESULT 35
US-09-914-259-61
; Sequence 61, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-259-61

Query Match      56.6%; Score 34.5; DB 4; Length 227;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY      1 RAQEK--VLQKLGKA 13
DB      44 RAQERLATVLOKLEEA 59

RESULT 36
US-09-107-532A-3730
; Sequence 3730, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:

US-09-107-532A-6086
; Sequence 6086, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3730:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...215
SEQUENCE DESCRIPTION: SEQ ID NO: 3730:
US-09-107-532A-3730.

Query Match      55.7%; Score 34; DB 4; Length 215;
Best Local Similarity 77.8%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 RAQEKVLQK 9
DB      205 RQEKVLKK 213

RESULT 37
US-09-107-532A-6086
; Sequence 6086, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
```



```
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...589
; SEQUENCE DESCRIPTION: SEQ ID NO: 6086:
US-09-107-532A-6086
Query Match 55.7%; Score 34; DB 4; Length 589;
Best Local Similarity 77.8%; Pred. No. 5.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QEKVLQKLG 11
DB 497 QEKVLETLG 505

RESULT 38
US-08-808-277A-3
; Sequence 3, Application US/08808277A
; Patent No. 5998374
; GENERAL INFORMATION:
; APPLICANT: LEHRER, ROBERT I.
; APPLICANT: LEHRER, ROBERT I.
; APPLICANT: LEE, IN-HEE
; APPLICANT: ZHAO, CHENGQUAN
; TITLE OF INVENTION: CLAVASPIRINS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,277A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 220002057300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-808-277A-3
Query Match 54.1%; Score 33; DB 2; Length 80;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLG 12
DB 27 EEKVQFLGK 36

RESULT 39
US-08-746-160-3
; Sequence 3, Application US/08746160
```

```
; Patent No. 6010876
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia I.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Lee, In-Hee
; TITLE OF INVENTION: CLAVANINS
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,160
; FILING DATE: 06-NOV-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20563.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-746-160-3
Query Match 54.1%; Score 33; DB 3; Length 80;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLG 12
DB 27 EEKVQFLGK 36

RESULT 40
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 6194555
; GENERAL INFORMATION:
; APPLICANT: Stocco, Douglas M.
; APPLICANT: Clark, Dr. Barbara J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
; TITLE OF INVENTION: OF STEROIDOGENESIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
; STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
; CITY: Austin
; STATE: TX
; COUNTRY: U.S.A.
; ZIP: 78701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,254
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/538,960
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mayfield, Denise L.
; REGISTRATION NUMBER: 33,732
; REFERENCE/DOCKET NUMBER: 43375.0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/499-6200
; TELEFAX: 512/499-6290
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-659-254-2

Query Match 54.1%; Score 33; DB 3; Length 276;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KVLQKLGK 12
Db 154 KVLQRIKG 161

RESULT 41
US-08-538-960-2
; Sequence 2, Application US/08538960
; Patent No. 5872230
; GENERAL INFORMATION:
; APPLICANT: Stocco, Douglas M.
; APPLICANT: Clark, Barbara J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: REGULATION OF STEROIDOGENESIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
; STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
; CITY: Austin
; STATE: TX
; COUNTRY: U.S.A.
; ZIP: 78701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,960
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mayfield, Denise L.
; REGISTRATION NUMBER: 33,732
; REFERENCE/DOCKET NUMBER: 43375.0002/DLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/499-6200
; TELEFAX: 512/499-6290
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 284 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-538-960-2

Query Match 54.1%; Score 33; DB 2; Length 284;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KVLQKLGK 12
Db 154 KVLQRIKG 161

RESULT 42
US-09-134-001C-4230
; Sequence 4230, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4230
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4230

Query Match 54.1%; Score 33; DB 4; Length 296;
Best Local Similarity 46.2%; Pred. No. 4.3e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RAQKVLQKLGKA 13
Db 138 RDEKLIQSVSKA 150

RESULT 43
US-08-455-910-13
; Sequence 13, Application US/08455910
; Patent No. 6221640
; GENERAL INFORMATION:
; APPLICANT: Tao, Jianshi
; APPLICANT: Sassanfar, Mandana
; APPLICANT: Gallant, Paul L.
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Avruch, Anthony S.
; APPLICANT: Yu, Russell V.
; APPLICANT: Nair, Shamila
; TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-TRNA SYNTHETASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/855,910
```

/ FILING DATE: 14-MAY-1997  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Brook, David E.  
/ REGISTRATION NUMBER: 22,592  
/ REFERENCE/DOCKET NUMBER: CPI95-08  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (781) 861-6240  
/ TELEFAX: (781) 861-9540  
/ INFORMATION FOR SEQ ID NO: 13:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 423 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ US-08-855-910-13

Query Match 54.1%; Score 33; DB 3; Length 423;  
Best Local Similarity 75.0%; Pred. No. 6e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 EKVQLKLG 11  
Db 307 EEILQKLG 314  
:::|||||

RESULT 44  
US-09-134-000C-5142  
/ Sequence 5142, Application US/09134000C  
/ Patent No. 6617156  
/ GENERAL INFORMATION:  
/ APPLICANT: Lynn Doucette-Stamm et al  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
/ FILE REFERENCE: 032796-032  
/ CURRENT APPLICATION NUMBER: US/09/134,000C  
/ CURRENT FILING DATE: 1998-08-13  
/ PRIOR FILING DATE: 1997-08-15  
/ NUMBER OF SEQ ID NOS: 612  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO 5142  
/ LENGTH: 427  
/ TYPE: PRT  
/ ORGANISM: Enterococcus faecalis  
/ US-09-134-000C-5142

Query Match 54.1%; Score 33; DB 4; Length 427;  
Best Local Similarity 75.0%; Pred. No. 6e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 EKVQLKLG 11  
Db 311 EEILQKLG 318  
:::|||||

RESULT 45  
US-08-477-451-18  
/ Sequence 18, Application US/08477451  
/ Patent No. 5928865  
/ GENERAL INFORMATION:  
/ APPLICANT: Covacci, Antonello  
/ TITLE OF INVENTION: Helicobacter Pylori Cagi Region  
/ NUMBER OF SEQUENCES: 46  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Chiron Corporation  
/ STREET: 4560 Horton Street  
/ CITY: Emeryville  
/ STATE: CA  
/ COUNTRY: USA  
/ ZIP: 94608-2916  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/477,451  
/ FILING DATE: 07-JUN-1995  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: McClung, Barbara G.  
/ REGISTRATION NUMBER: 33,113  
/ REFERENCE/DOCKET NUMBER: 0335.002  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 510-601-2708  
/ TELEFAX: 510-655-3542  
/ INFORMATION FOR SEQ ID NO: 18:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 464 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ US-08-477-451-18

Query Match 54.1%; Score 33; DB 2; Length 464;  
Best Local Similarity 41.7%; Pred. No. 6.5e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAEKVKLQKLG 12  
Db 269 KATKILSIGK 280  
:::|||||

RESULT 46  
US-08-313-075A-50  
/ Sequence 50, Application US/08313075A  
/ Patent No. 5639870  
/ GENERAL INFORMATION:  
/ APPLICANT: Holton, Timothy A.  
/ APPLICANT: Cornish, Edwina C.  
/ APPLICANT: Tanaka, Yoshikazu  
/ TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID  
/ TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR  
/ NUMBER OF SEQUENCES: 58  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Scully, Scott, Murphy & Presser  
/ STREET: 400 Garden City Plaza  
/ CITY: Garden City  
/ STATE: New York  
/ COUNTRY: U.S.A.  
/ ZIP: 11530

COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/313,075A  
/ FILING DATE: 30-NOV-1994  
/ CLASSIFICATION: 800  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: AU PL 1538/92  
/ FILING DATE: 27-MAR-1992  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: AU PL 6698/93  
/ FILING DATE: 07-JAN-1993  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: AU PCT/AU93/00127  
/ FILING DATE: 25-MAR-1993  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Digiglio, Frank S.  
/ REGISTRATION NUMBER: 31,346  
/ REFERENCE/DOCKET NUMBER: 9433  
/ TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR 50:  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 496 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-313-075A-50

Query Match 54.1%; Score 33; DB 1; Length 496;  
Best Local Similarity 50.0%; Pred. No. 7e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
Db 325 KAQELAQVIGK 336

## RESULT 47

US-09-519-232-8  
Sequence 8, Application US/09519232  
Patent No. 6528702  
GENERAL INFORMATION:  
APPLICANT: Salmeron, John  
APPLICANT: Weisio, Laura  
APPLICANT: Willits, Michael  
APPLICANT: Mengiste, Tesfaye  
TITLE OF INVENTION: NOVEL PLANT GENES AND USES THEREOF  
FILE REFERENCE: S-30857A/RTP2095  
CURRENT APPLICATION NUMBER: US/09/519,232  
CURRENT FILING DATE: 2000-03-06  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 8  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-519-232-8

Query Match 54.1%; Score 33; DB 4; Length 586;  
Best Local Similarity 41.7%; Pred. No. 8.1e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
Db 254 KISEKLERIGK 265

## RESULT 48

US-09-107-532A-6359  
Sequence 6359, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6359:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 822 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (3) LOCATION 1...822  
SEQUENCE DESCRIPTION: SEQ ID NO: 6359:  
US-09-107-532A-6359

Query Match 54.1%; Score 33; DB 4; Length 822;  
Best Local Similarity 60.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AQEKVLQKLG 11  
Db 612 AQEILQKFG 621

## RESULT 49

US-09-543-681A-5053  
Sequence 5053, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 5053  
LENGTH: 826  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-5053

Query Match 54.1%; Score 33; DB 4; Length 826;  
Best Local Similarity 77.8%; Pred. No. 1.1e+03;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AQEKVLQKLG 10  
Db 193 AQSRVLQKL 201

## RESULT 50

US-09-540-236-3299  
Sequence 3299, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 3299

; LENGTH: 1105

; TYPE: PRT

; ORGANISM: M.catarrahals

US-09-540-236-3299

Query Match 54.1%; Score 33; DB 4; Length 1105;

Best Local Similarity 60.0%; Pred. No. 1.5e+03;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QEKVLQKLGK 12

:|:|:|:|

Db 117 EQKVEQKIGK 126

Search completed: March 4, 2004, 17:47:37

Job time : 15.4194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:31:50 ; Search time 58.6452 Seconds

(without alignments)  
86.723 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_138\_155

Perfect score: 95

Sequence: 1 RIAXGRKLVVDVSARHH 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : A Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003s:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	130	4	ABG22449 Novel hum
2	95	100.0	130	4	ABG20881 Novel hum
3	95	100.0	130	4	ABG21407 Novel hum
4	95	100.0	130	4	ABG20318 Novel hum
5	95	100.0	252	4	AAW62101 Bar domai
6	95	100.0	404	2	AAW05392 Human SH3
7	95	100.0	414	4	ABG13005 Novel hum
8	95	100.0	425	4	ABG22456 Novel hum
9	95	100.0	434	2	AAW05391 Mouse SH3
10	95	100.0	431	2	AAW06602 Human box
11	95	100.0	451	2	AAW47295 Human Bin
12	95	100.0	451	2	AAW36881 Human Bin
13	95	100.0	451	2	AAW94504 Human Bin
14	95	100.0	453	6	ABR95631 Human CGD
15	95	100.0	482	6	ABU99771 Protein d
16	95	100.0	564	4	AAW62100 Human bri
17	95	100.0	572	4	ABG21411 Novel hum
18	95	100.0	588	7	ADD45951 Rat Prote
19	95	100.0	588	7	ADG60848 Rat Prote
20	95	100.0	588	7	ADG60840 Rat Prote
21	95	100.0	588	7	ADG60844 Rat Prote
22	95	100.0	588	7	ADG60852 Rat Prote
23	95	100.0	593	7	ADG60842 Human Pro
24	95	100.0	593	7	ADD45953 Human Pro
25	95	100.0	593	7	ADG60850 Human Pro

26	95	100.0	593	7	ADE60846	Human Pro
27	95	100.0	593	7	ADB60854	Human Pro
28	95	100.0	594	4	ABG20887	Novel hum
29	95	100.0	594	4	ABG20324	Novel hum
30	95	100.0	594	4	ABG22457	Novel hum
31	95	100.0	960	4	ABG20886	Novel hum
32	95	100.0	960	4	ABG20323	Novel hum
33	89	93.7	683	7	ADD44887	Rat Prote
34	89	93.7	695	5	AAE22091	Human amp
35	89	93.7	695	5	AAU75110	Synaptic
36	89	93.7	695	7	ADD44889	Human Pro
37	62	65.3	602	4	ABB63948	Drosophil
38	49	51.6	276	6	ADA48708	Rice prot
39	47	49.5	424	5	ABB91883	Herbicida
40	47	49.5	425	4	AAW48428	Protein A
41	47	49.5	468	3	AAW41757	Arabidops
42	47	49.5	472	3	AAW41756	Arabidops
43	47	49.5	587	3	AAW41755	Arabidops
44	47	49.5	765	6	ABU24499	Protein e
45	45.5	47.9	521	5	ABP55425	Mouse pol
46	45.5	47.9	578	6	ABU96724	Human nuc
47	45.5	47.9	596	4	AAW95815	Human pro
48	45.5	47.9	742	7	ADB64808	Human pro
49	45.5	47.9	763	6	ABU00410	Human nov
50	44	46.3	161	4	AAU63157	Propionib
51	44	46.3	161	6	ABW59676	Propionib
52	44	46.3	297	6	ABU23306	Protein e
53	43.5	45.8	282	6	ABU47137	Protein e
54	43.5	45.8	282	6	ABU47874	Protein e
55	43	45.3	848	7	ADE54727	Rat Prote
56	42.5	44.7	304	6	ABU45447	Protein e
57	42	44.2	130	5	ABP28276	Streptoco
58	42	44.2	463	6	ABP57491	Mycobacte
59	42	44.2	463	6	ABU36480	Protein e
60	42	44.2	463	6	ABU34382	Protein e
61	42	44.2	634	3	ABU33599	Modified
62	42	44.2	634	4	AAU04709	Modified
63	42	44.2	634	4	AAU05034	Modified
64	42	44.2	712	6	ABU29750	Protein e
65	42	44.2	721	7	ADC97262	E. faeciu
66	42	44.2	903	2	AAW28713	Detargete
67	42	44.2	906	2	AAW28712	Mutant ye
68	42	44.2	906	2	AAW28702	Yeast mul
69	42	44.2	906	4	AAW20182	C. tropic
70	42	44.2	906	4	AAW20181	C. tropica
71	42	44.2	906	4	AAW20180	Candida t
72	42	44.2	906	4	AAW20183	C. tropic
73	41.5	43.7	731	2	AAW60822	Melon pro
74	41.5	43.7	763	2	AAW24094	Mouse EPC
75	41	43.2	90	4	AAW65249	Snowshoe
76	41	43.2	196	4	ABW69788	Drosophil
77	41	43.2	237	5	ABU05849	M. tuberc
78	41	43.2	385	4	ABG11090	Novel hum
79	41	43.2	385	4	ABG16959	Novel hum
80	41	43.2	771	6	ABJ25622	Aspergill
81	41	43.2	905	6	ABJ26222	Novel hum
82	40.5	42.6	187	4	ABG02458	Novel hum
83	40.5	42.6	274	4	ABG17857	Novel hum
84	40.5	42.6	298	7	ADC94099	E. faeciu
85	40.5	42.6	306	6	ABU15207	Protein e
86	40	42.1	124	4	ABG08236	Novel hum
87	40	42.1	142	4	ABG08236	Proteinib
88	40	42.1	142	6	ABW64337	Proteinib
89	40	42.1	171	4	AAW64390	Amino aci
90	40	42.1	171	4	AAW40219	Human pol
91	40	42.1	216	4	AAW93722	Human pro
92	40	42.1	244	4	ABG08238	Novel hum
93	40	42.1	259	4	AAW42005	Human pol
94	40	42.1	287	6	AAW11735	Human MDD
95	40	42.1	290	6	ABU70917	Human adi
96	40	42.1	294	3	ABW07857	Amino aci
97	40	42.1	298	6	ABP97410	Thermoana
98	40	42.1	321	4	AAU54500	Propionib

99 40 42.1 321 6 ABM51019  
100 40 42.1 344 4 AAM33186  
101 40 42.1 406 3 AAY94209  
102 40 42.1 406 7 AAB79885  
103 40 42.1 414 3 AAB07856  
104 40 42.1 414 3 AAB58368  
105 40 42.1 414 5 ABP41711  
106 40 42.1 436 6 ABU99782  
107 40 42.1 440 4 ABB72014  
108 40 42.1 460 3 AAB33157  
109 40 42.1 625 5 ABU10986  
110 40 42.1 634 4 ABG15117  
111 40 42.1 701 6 ABU4429  
112 40 42.1 735 4 ABB69024  
113 40 42.1 752 6 ABR52722  
114 40 42.1 1059 4 ABB59330  
115 40 42.1 1059 4 ABB67407  
116 40 42.1 1059 4 ABB67408  
117 39 41.1 89 6 ABU23366  
118 39 41.1 89 6 ABU17088  
119 39 41.1 95 6 ADA36851  
120 39 41.1 147 4 AAM40760  
121 39 41.1 210 4 ABB08543  
122 39 41.1 245 5 ABP27803  
123 39 41.1 245 6 ABU46473  
124 39 41.1 259 3 AAG13065  
125 39 41.1 265 3 AAG13064  
126 39 41.1 269 3 AAG13063  
127 39 41.1 335 3 AAY59252  
128 39 41.1 335 3 AAB10485  
129 39 41.1 335 3 AAY90285  
130 39 41.1 335 4 AAM38974  
131 39 41.1 392 3 AAG47417  
132 39 41.1 444 6 ABU34078  
133 39 41.1 450 4 ABB62490  
134 39 41.1 460 4 ABB58088  
135 39 41.1 471 6 ABJ25422  
136 39 41.1 481 6 ABU25218  
137 39 41.1 493 3 AAG47416  
138 39 41.1 512 4 ABB60708  
139 39 41.1 515 3 AAG17527  
140 39 41.1 516 3 AAG47415  
141 39 41.1 523 3 AAG17526  
142 39 41.1 525 3 AAG17525  
143 39 41.1 565 6 ABB26022  
144 39 41.1 626 6 ABU52572  
145 39 41.1 708 4 ABB70030  
146 39 41.1 708 4 ABB67226  
147 39 41.1 853 6 ABU25701  
148 39 41.1 892 6 ABP96857  
149 39 41.1 981 3 AAY97012  
150 38.5 40.5 52 4 AAU45270

ALIGNMENTS

RESULT 1  
ABG22449  
ID ABG22449 standard; protein; 130 AA.  
XX AC  
XX ABG22449;  
XX

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #22440.

XX Human; chromosome mapping; gene therapy; forensic;  
KW Food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

ABM51019 Propionib  
AAM33186 Human pol  
AAY94209 Human TRA  
AAB79885 Human put  
AAB07856 Amino aci  
AAB58368 Lung canc  
ABP41711 Human ova  
ABU99782 Novel hum  
ABB72014 Drosophil  
AAB33157 Human ORF  
ABJ10986 Yeast sel  
ABG15117 Novel hum  
ABU4429 Protein e  
ABB69024 Drosophil  
ABR52722 Protein s  
ABB59330 Drosophil  
ABB67407 Drosophil  
ABB67408 Drosophil  
ABU23366 Protein e  
ABU17088 Protein e  
ADA36851 Acinetoba  
AAM40760 Human pol  
ABB08543 Novel hum  
ABP27803 Streptoco  
AAB46473 Protein e  
AAG13065 Arabidops  
AAG13064 Arabidops  
AAG13063 Arabidops  
AAY59252 Human gly  
AAB10485 Human MPR  
AAY90285 Human pep  
AAM38974 Human pol  
AAG47417 Arabidops  
ABU34078 Protein e  
ABB62490 Drosophil  
ABB58088 Drosophil  
ABJ25422 Aspergill  
ABU25218 Protein e  
AAG47416 Arabidops  
ABB60708 Drosophil  
AAG17527 Arabidops  
AAG47415 Arabidops  
AAG17526 Arabidops  
AAG17525 Arabidops  
ABJ26022 Aspergill  
ABU52572 Peanut Ar  
ABB70030 Drosophil  
ABB67226 Drosophil  
ABU25701 Protein e  
ABP96857 Escherich  
AAY97012 S. cerevi  
AAU45270 Propionib

XX 11-OCT-2001.  
XX 30-MAR-2001; 2001WO-US008631.  
XX 31-MAR-2000; 2000US-00540217.  
XX 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX N-PSDB; AAS86636.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX Claim 20; SEQ ID NO 52808; 103pp; English.  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
XX patent did not appear in the printed specification. Note: The sequence data for this  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 130 AA;  
XX Query Match 100.0%; Score 95; DB 4; Length 130;  
XX Best Local Similarity 100.0%; Pred. No. 2.6e-08;  
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RIARGRKLVYDSARHH 18  
Db 111 RIARGRKLVYDSARHH 128  
RESULT 2  
ABG20881  
ID ABG20881 standard; protein; 130 AA.  
XX AC  
XX ABG20881;  
XX  
XX 13-FEB-2002 (first entry)  
XX Novel human diagnostic protein #20872.  
XX Human; chromosome mapping; gene therapy; forensic;  
KW Food supplement; medical imaging; diagnostic; genetic disorder.  
XX Homo sapiens.  
XX WO200175067-A2.  
XX 11-OCT-2001.  
PD

XX 30-MAR-2001; 2001WO-US008631.  
PF  
XX 31-MAR-2000; 2000US-00540217.  
PR  
XX 23-AUG-2000; 2000US-00649167.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR  
XX N-PSDB; AAS85068.  
DR  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
PT  
XX  
XX Claim 20; SEQ ID NO 51240; 103pp; English.  
PS  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
XX amino acid sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 130 AA;  
Query Match 100.0%; Score 95; DB 4; Length 130;  
Best Local Similarity 100.0%; Pred. No. 2.6e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 RIAKGRKLVYDSARHH 18  
Db 111 RIAKGRKLVYDSARHH 128  
RESULT 3  
ABG21407  
ID ABG21407 standard; protein; 130 AA.  
XX  
XX AC ABG21407;  
XX  
XX DT 18-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #21398.  
XX  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX  
XX PF

XX 31-MAR-2000; 2000US-00540217.  
PR  
XX 23-AUG-2000; 2000US-00649167.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR  
XX N-PSDB; AAS85594.  
DR  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
PT  
XX  
XX Claim 20; SEQ ID NO 51766; 103pp; English.  
PS  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
XX amino acid sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 130 AA;  
Query Match 100.0%; Score 95; DB 4; Length 130;  
Best Local Similarity 100.0%; Pred. No. 2.6e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 RIAKGRKLVYDSARHH 18  
Db 111 RIAKGRKLVYDSARHH 128  
RESULT 4  
ABG20318  
ID ABG20318 standard; protein; 130 AA.  
XX  
XX AC ABG20318;  
XX  
XX DT 13-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #20309.  
XX  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX  
XX PF 31-MAR-2000; 2000US-00540217.  
XX  
XX PF



PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS84505.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 20; SEQ ID NO 50677; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 130 AA;

Query Match 100.0%; Score 95; DB 4; Length 130;

Best Local Similarity 100.0%; Pred. No. 2.6e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRGKLVYDSARHH 18

Db 111 RIAGRGKLVYDSARHH 128

RESULT 5

AAB62101

ID AAB62101 standard; protein; 252 AA.

XX AAB62101;

XX 29-MAY-2001 (first entry)

XX Bar domain of Bin1 protein.

XX Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
KW hyperplastic disease; cytostatic; cell growth regulator; Bin1;  
KW chromosome 4q22.1.

XX Homo sapiens.

XX WO200116158-A2.

XX 08-MAR-2001.

XX 30-AUG-2000; 2000WO-US023723.

XX 31-AUG-1999; 99US-0151554P.

(WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Prendergast GC, Ge K;

XX WPI; 2001-235087/24.

XX New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
PT useful for regulating cell growth, and for diagnosing or treating  
PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
PT or hepatocarcinoma.

XX Example 3; Fig 2; 62pp; English.

XX The invention provides a human bridging integrator-2 (Bin2) protein. The  
CC protein can be expressed by standard recombinant methodology. The Bin2  
CC proteins or peptides are useful in regulating cell growth, cell survival,  
CC differentiation, endocytosis and actin organization. These peptides or  
CC proteins are also useful for diagnosing or treating conditions associated  
CC with inappropriate expression of Bin2, e.g. liver cancer, breast cancer,  
CC hepatocarcinoma, myeloid and lymphoid leukemias or blood disorders. The  
CC peptides, nucleic acid sequences or anti-Bin2 antibodies are useful for  
CC diagnosing inappropriate expression of Bin2 Bin2 is also useful for  
CC treating disorders associated with excessive Bin1 levels, e.g. liver,  
CC colorectal, prostate or breast cancers, epithelia cell cancers, melanoma,  
CC or hyperplastic disease states. The present sequence represents the Bar  
CC domain of Bin1 used in homology studies with Bin2 protein

XX Sequence 252 AA;

Query Match 100.0%; Score 95; DB 4; Length 252;

Best Local Similarity 100.0%; Pred. No. 5.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRGKLVYDSARHH 18

Db 139 RIAGRGKLVYDSARHH 156

RESULT 6

AAB05392

ID AAB05392 standard; protein; 404 AA.

XX AAB05392;

XX 19-FEB-1998 (first entry)

XX Human SH3P9 protein.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
KW cellular signalling element; cellular structural element; malignancy;  
KW protein identification; functional domain; protein screening;  
KW cellular signal transduction process.

XX Homo sapiens.

XX WO9631625-A1.

XX 10-OCT-1996.

XX 04-APR-1996; 96WO-US004454.

XX 07-APR-1995; 95US-00417872.

XX 03-APR-1996; 96US-00630915.

XX (CYTO-) CYTOGEN CORP.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX WPI; 1996-465045/46.

XX N-ESDB; AAT39792.

XX Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,  
 PT regardless of sequence homology.

PS Claim 54; Fig 37; 174pp; English.

XX AA05386-W05403 represent novel human and mouse Src-homology region 3  
 CC (SH3) domain containing proteins that can be used in the method of the  
 CC invention. SH3 domain containing proteins play a role in signalling and  
 CC structural elements of cells. The method of the invention is for  
 CC identifying polypeptides containing functional domains of interest  
 CC (especially SH3 domains). The method comprises contacting a multivalent  
 CC recognition unit (RU) complex with a number of peptides and identifying  
 CC polypeptides having a selective binding affinity for the RU complex. The  
 CC method is based on functional similarities and does not rely on sequence  
 CC similarities. Prior methods only gave limited success for identifying  
 CC proteins which contain an SH3 domain due to the minimal sequence homology  
 CC among known SH3 proteins. It has been found that small peptide RUs in  
 CC multivalent form have reduced specificity for a given functional domain  
 CC compared to monomer RUs. Multivalent RU complexes are particularly suited  
 CC to screening for polypeptides containing functional domains that are  
 CC similar to, but not identical in sequence to, the original target  
 CC functional domain. The new method enables proteins having a common  
 CC function to be identified. Identification of novel SH3 proteins will be  
 CC useful for a better understanding of cell growth, malignancy, signal  
 CC transduction processes, etc. New candidate drugs can be identified, and  
 CC their specificities (e.g. pharmacological activities) can be assessed  
 CC using the method of the invention

XX Sequence 404 AA;

Query Match 100.0%; Score 95; DB 2; Length 404;

Best Local Similarity 100.0%; Pred. No. 9.1e-08; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 0

QY 1 RIAKGRKLVYDSARHH 18

Db 104 RIAKGRKLVYDSARHH 121

RESULT 7

ABG13005

ID ABG13005 standard; protein; 414 AA.

AC ABG13005;

XX 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #12996.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS77192.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 43364; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 414 AA;

Query Match 100.0%; Score 95; DB 4; Length 414;

Best Local Similarity 100.0%; Pred. No. 9.3e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18

Db 138 RIAKGRKLVYDSARHH 155

RESULT 8

ABG22456

ID ABG22456 standard; protein; 425 AA.

XX ABG22456;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #22447.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS86643.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

PS Claim 20; SEQ ID NO 52815; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed CC genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC useful for generating antibodies against it, detecting or quantitating a CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activities. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic CC amino acid sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published\_pct\_sequences XX

SQ Sequence 425 AA;

Query Match 100.0%; Score 95; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAKRGKLVYDSARHH 18

Db 97 RIAKRGKLVYDSARHH 114

RESULT 9

AAW05391

ID AAW05391 standard; protein; 434 AA.

XX AC AAW05391;

DT 18-FEB-1998 (first entry)

XX DE Mouse SH3P9 protein.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
cellular signalling element; cellular structural element; malignancy;  
protein identification; functional domain; protein screening;  
cellular signal transduction process.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

FT Misc-difference 433 /note= "encoded by CTA"

XX PN W09631625-A1.

XX PD 10-OCT-1996.

XX PF 04-APR-1996; 96WO-US004454.

XX PR 07-APR-1995; 95US-00417872.

XX PR 03-APR-1996; 96US-00630915.

XX PA (CVTO-) CVTOGEN CORP.

XX PA (UYNC-) UNIV NORTH CAROLINA.

XX PI Sparks AB, Hoffman N, Kay EK, Fowlkes DM, McConnell SJ;

XX DR WPI; 1996-465045/46.

XX DR N-PSDB; AAT39791.

XX

PT Identifying polypeptide(s) having specific functional domain (esp. SH3  
PT domain) - comprises detecting selective binding to recognition unit,  
PT regardless of sequence homology.

XX Claim 54; Fig 35; 174pp; English.

XX AAW05386-W05403 represent novel human and mouse Src-homology region 3  
CC (SH3) domain containing proteins that can be used in the method of the  
CC invention. SH3 domain containing proteins play a role in signalling and  
CC structural elements of cells. The method of the invention is for  
CC identifying polypeptides containing functional domains of interest  
CC (especially SH3 domains). The method comprises contacting a multivalent  
CC recognition unit (RU) complex with a number of peptides and identifying  
CC polypeptides having a selective binding affinity for the RU complex. The  
CC method is based on functional similarities and does not rely on sequence  
CC similarities. Prior methods only gave limited success for identifying  
CC proteins which contain an SH3 domain due to the minimal sequence homology  
CC among known SH3 proteins. It has been found that small peptide RUS in  
CC multivalent form have reduced specificity for a given functional domain  
CC compared to monomer RUS. Multivalent RU complexes are particularly suited  
CC to screening for polypeptides containing functional domains that are  
CC similar to, but not identical in sequence to, the original target  
CC functional domain. The new method enables proteins having a common  
CC function to be identified. Identification of novel SH3 proteins will be  
CC useful for a better understanding of cell growth, malignancy, signal  
CC transduction processes, etc. New candidate drugs can be identified, and  
CC their specificities (e.g. pharmacological activities) can be assessed  
CC using the method of the invention XX

SQ Sequence 434 AA;

Query Match 100.0%; Score 95; DB 2; Length 434;

Best Local Similarity 100.0%; Pred. No. 9.8e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAKRGKLVYDSARHH 18

Db 139 RIAKRGKLVYDSARHH 156

RESULT 10

AAW06602

ID AAW06602 standard; protein; 451 AA.

XX AC AAW06602;

XX DT 07-FEB-1997 (first entry)

XX DE Human box-dependent myc-interacting protein (BIN1).

XX Box-dependent myc-interacting protein; BIN1; MIP-99; oncogene;  
XX oncoprotein; breast cancer; liver cancer; apoptosis;  
XX tumour suppressor protein; hyperplasia; diagnosis; therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 1. .222 /note= "polypeptide useful in constructs of the invention

FT Domain (Claim 9)"

FT Domain 190. .250 /note= "polypeptide useful in constructs of the

FT Domain invention (Claim 9)"

FT Domain 223. .251 /note= "polypeptide useful in constructs of the invention

FT Domain (Claim 9)"

FT Domain 252. .265 /label= Nuclear localisation domain

FT Domain (Claim 9)"

FT Domain 263. .397 /note= "polypeptide useful in constructs of the invention

FT Domain (Claim 9)"

FT Domain 263. .397 /note= "polypeptide useful in constructs of the invention

FT Domain (Claim 9)"

FT Domain 263. .397 /note= "polypeptide useful in constructs of the invention

FT Domain (Claim 9)"

```

FT Domain 270. .383
FT /label= Myc-binding_domain
FT /note= "polypeptide useful in constructs of the invention
FT (claim 9)"
FT 278. .451
FT /label= SH3
FT /note= "SH3 domain useful in constructs of the invention
FT (claim 9); this domain may be useful in blocking
FT apoptosis"
XX
XX WO9634627-A1.
XX
XX 07-NOV-1996.
XX
XX 02-MAY-1996; 96WO-US006231.
XX
XX 05-MAY-1995; 95US-00435454.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Prendergast GC, Sakamuro D;
XX
XX WPI; 1996-505902/50.
XX N-PSDB; AAT43573.
XX
XX Mammalian box-dependent myc-interacting protein Binl coding sequences -
XX used in the diagnosis and treatment of c-myc associated cancers and
XX hyperplastic disease and Binl deficiency.
XX
XX Claim 3; Page 44-46; 79pp; English.
XX
XX Human box-dependent myc-interacting protein Binl (AAW06602) is a putative
XX tumour suppressor protein. It inhibits Myc-dependent malignant cell
XX transformation, and expression of Binl mRNA is lost in human liver and
XX breast carcinoma cell lines. The amino acid sequence of human Binl was
XX deduced from a cDNA clone (AAT43573) isolated from a human skeletal
XX muscle cDNA library. Recombinant Binl or its polypeptides can be produced
XX in host cells for use in the detection, diagnosis and treatment of Binl
XX deficiency diseases and cancers or hyperplastic diseases involving the c-
XX myc oncogene
XX
XX Sequence 451 AA;
XX
XX Query Match 100.0%; Score 95; DB 2; Length 451;
XX Best Local Similarity 100.0%; Pred. NO. 1e-07;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RIAGRGKLVVDYDSARHH 18
XX |||||
XX 136 RIAGRGKLVVDYDSARHH 153
XX
XX RESULT 11
XX ID AAW47295
XX AC AAW47295 standard; protein; 451 AA.
XX AC AAW47295;
XX
XX DT 01-JUN-1998 (first entry)
XX
XX DE Human Binl.
XX
XX KW Human; Box-dependent myc-interacting protein; Binl; treatment; cancer;
XX hyperplastic disease; neurodegenerative disease; antibody; diagnosis.
XX
XX OS Homo sapiens.
XX
XX PS US5723581-A.
XX
XX PD 03-MAR-1998.
XX
XX PF 24-MAY-1996; 96US-00652972.
XX

```

```

PR 05-MAY-1995; 95US-00435454.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Sakamuro D, Prendergast GC;
XX
XX WPI; 1998-178541/16.
XX N-PSDB; AAV15692.
XX
XX Box-dependent myc-interacting protein Binl - useful for, e.g. treating
XX diseases associated with inadequate Binl levels such as colon cancer.
XX
XX Claim 1; Col 25-30; 32pp; English.
XX
XX The present sequence is human Box-dependent myc-interacting protein Binl,
XX which can be used to treat diseases associated with inadequate Binl
XX levels, e.g. liver, colon or breast cancer or hyperplastic diseases, or
XX diseases where Binl activity needs to be inhibited, e.g.
XX neurodegenerative diseases, or to raise antibodies for diagnostic
XX purposes
XX
XX Sequence 451 AA;
XX
XX Query Match 100.0%; Score 95; DB 2; Length 451;
XX Best Local Similarity 100.0%; Pred. NO. 1e-07;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 RIAGRGKLVVDYDSARHH 18
XX |||||
XX 136 RIAGRGKLVVDYDSARHH 153
XX
XX RESULT 12
XX ID AAB36881
XX AC AAB36881 standard; protein; 451 AA.
XX AC AAB36881;
XX
XX DT 22-FEB-2001 (first entry)
XX
XX DE Human Binl protein.
XX
XX KW Human; BIN1-associated U1-specific; Bau; cancer; gene therapy.
XX
XX OS Homo sapiens.
XX
XX PN US6140465-A.
XX
XX PD 31-OCT-2000.
XX
XX PF 25-JUN-1999; 99US-00344889.
XX
XX PR 29-AUG-1996; 96US-0025482P.
XX PR 27-AUG-1997; 97US-00919145.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Prendergast GC;
XX
XX WPI; 1998-207038/18.
XX N-PSDB; AAC68498.
XX
XX New isolated BIN1-associated U1-specific protein - which acts as a tumour
XX suppressor, used to develop products for treating e.g. cancers, hyper-
XX plastic disease states or degenerative diseases.
XX
XX Disclosure; Fig 3; 22pp; English.
XX
XX The present invention relates to mammalian BIN1-associated U1-specific
XX (Bau) protein. The Bau protein is useful in the detection, diagnosis and
XX treatment of cancers or other disorders associated with inappropriate
XX BIN1 levels and/or deregulation, deficiency or amplification of the c-Myc
XX oncogenes. The proteins are also useful as antigens for the development
XX

```

CC of anti-Bau antiserum and antibodies to Bau, or to a desired fragment of  
 CC the Bau protein, as diagnostic reagents, in gene therapy, and in  
 CC screening and developing chemical compounds or proteins which may be used  
 CC for the treatment of cancers characterized by Bau or BIN1, which regulate  
 CC inappropriate MYC levels  
 XX SQ Sequence 451 AA;

Query Match 100.0%; Score 95; DB 2; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 1e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 |||||  
 Db 136 RIAGRGKLVYDSARHH 153

RESULT 13  
 ID AAW94504 standard; protein; 451 AA.

XX AC AAW94504;  
 XX DT 22-APR-1999 (first entry)  
 XX DE Human Bin1 protein.

XX KW Bin1; brain-specific; box-dependent myc-interacting protein; cancer;  
 XX diagnosis; hyperplastic disease; tumour suppressor; gene therapy;  
 XX benign prostatic hypertrophy; neurodegeneration.

XX OS Homo sapiens.

XX PN WO9855151-A1.

XX PD 10-DEC-1998.

XX PF 04-JUN-1998; 98WO-US011647.

XX PR 06-JUN-1997; 97US-00870126.

XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX PI Prendergast GC, Sakamuro D;

XX DR WPI; 1999-059881/05.

XX DR N-PSDB; AAX16322.

PT New splice variant exons of Box-dependent myc-interacting polypeptide -  
 PT associated with loss of tumour suppressor activity, used for diagnosis of  
 PT cancer and in gene therapy.

XX PS Example 1; Page 81-82; 133pp; English.

XX CC The present invention describes Bin1 (Box-dependent myc-interacting  
 CC protein 1) brain-specific alternative splice variants exon 12A, B, C and  
 CC D. Bin1 specific antibodies (when labelled) are used to detect cancers or  
 CC other hyperplastic conditions (e.g. benign prostatic hypertrophy)  
 CC associated with a deficit of normal Bin1 and/or aberrant forms of Bin1.  
 CC The same diseases can also be diagnosed at the nucleic acid level using  
 CC fragments of Bin1 nucleotide sequences in standard amplification and/or  
 CC hybridisation assays. Cancers that can be detected are carcinomas and  
 CC epithelial cell tumours, specifically of prostate, liver and colon/  
 CC rectum, also melanoma. Bin1 nucleotide sequences, when included in a  
 CC vector, and Bin1 proteins may also be used to treat these diseases, and  
 CC also degenerative conditions such as neurodegeneration. Bin1 proteins may  
 CC also be used to raise antibodies, and the nucleotide sequences can be  
 CC used to express the corresponding proteins. Also anti-idiotypic  
 CC antibodies can be used therapeutically, and more generally similar  
 CC materials can be used to treat or diagnose any condition involving  
 CC deregulation, defect or amplification of the c-myc oncogene. Bin1  
 CC nucleotide sequence, proteins and antibodies are also useful to screen  
 CC for agents that may be used to treat Bin1-related cancers. The present

CC sequence represents human Bin1 protein  
 XX SQ Sequence 451 AA;

Query Match 100.0%; Score 95; DB 2; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 1e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 |||||  
 Db 136 RIAGRGKLVYDSARHH 153

RESULT 14  
 ID ABR69631 standard; protein; 453 AA.

XX AC ABR69631;

XX DT 12-AUG-2003 (first entry)

XX DE Human CGDD-31 protein.

XX KW Human; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
 XX neuroprotective; cerebroprotective; anti-HIV; anti-allergic;  
 XX anti-inflammatory; gynaecological; cancer; atherosclerosis; epilepsy;  
 XX Huntington's disease; stroke; AIDS; allergy; placenta; reproductive;  
 XX CGDD; cell growth; cell differentiation; cell death.

XX OS Homo sapiens.

XX PN WO2003027263-A2.

XX PD 03-APR-2003.

XX PF 26-SEP-2002; 2002WO-US031095.

XX PR 28-SEP-2001; 2001US-0326389P.

XX PR 05-OCT-2001; 2001US-0327380P.

XX PR 03-OCT-2001; 2001US-0328186P.

XX PR 12-OCT-2001; 2001US-0329690P.

XX PR 26-OCT-2001; 2001US-0345384P.

XX PR 26-OCT-2001; 2001US-0348165P.

XX PR 02-NOV-2001; 2001US-0350219P.

XX PR 09-NOV-2001; 2001US-0344518P.

XX PR 09-NOV-2001; 2001US-0345143P.

XX PR 16-NOV-2001; 2001US-0332375P.

XX PR 03-DEC-2001; 2001US-0336908P.

XX PR 07-DEC-2001; 2001US-0340747P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Chawla NK;

XX PI Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Gorvad AE;

XX PI Griffin JA, Hafalia AJA, Ison CH, Kable AE, Kalafus DP;

XX PI Lehr-Mason PM, Lu DM, Marquis JP, Nguyen DB, Ramkumar J;

XX PI Richardson TM, Sapperstein SK, Swarnakar A, Tang YT, Tran UK;

XX PI Warren BA, Xu Y, Yao MG, Yue H, Yue H;

XX DR WPI; 2003-421159/39.

XX DR N-PSDB; ACC30608.

XX CC New human proteins associated with cell growth, differentiation, and  
 XX death (CGDD), useful for diagnosing, treating and preventing diseases or  
 XX conditions associated with the aberrant CGDD expression e.g. cancer,  
 XX AIDS, or epilepsy.

XX PS Claim 1; Page 281-282; 350pp; English.

XX CC The invention relates to an isolated polypeptide associated with cell  
 XX growth, differentiation and death (CGDD). Also disclosed are the  
 XX polynucleotides encoding the polypeptides. The polypeptides and  
 XX polynucleotides are useful in diagnosing, treating and preventing

CC diseases or conditions associated with the decreased expression or over  
 CC expression of CGPD. Such diseases include cell proliferative (e.g.  
 CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's  
 CC disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and  
 CC reproductive disorders, or disorders of the placenta. They are also  
 CC useful in assessing the effects of exogenous compounds on the expression  
 CC of nucleic acid and amino acid sequences of CGPD. The CGPD or its  
 CC fragments are useful in screening compounds for effectiveness as an  
 CC agonist or antagonist of the polypeptides, or in altering the expression  
 CC of the target polynucleotide and compounds that specifically bind to or  
 CC modulate the activity of the polypeptide. Microarrays consisting  
 CC of polynucleotides of the invention are useful in monitoring or measuring  
 CC protein-protein interactions, drug-target interactions, and gene  
 CC expression profiles. Sequences given in records AAR69601-AAR69657  
 CC represent CGPD polypeptides of the invention  
 XX  
 SQ Sequence 453 AA;

Query Match 100.0%; Score 95; DB 6; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 1e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 |||||  
 DB 139 RIAGRGKLVYDSARHH 156

RESULT 15  
 ABU89771  
 ID ABU89771 standard; protein; 482 AA.

AC ABU89771;  
 XX  
 DT 10-JUL-2003 (first entry)  
 XX  
 DE Protein differentially expressed in cardiovascular disease #65.  
 XX  
 KW Cardiovascular disease; atherosclerosis; ischaemia; angina pectoris;  
 KW myocardial infarction; cardiac; antiatherosclerotic; antianginal;  
 KW gene therapy; differential gene expression.  
 XX  
 OS Homo sapiens.

XX WO2003031650-A2.  
 PN  
 XX 17-APR-2003.  
 PD

XX 02-OCT-2002; 2002WO-EP011034.  
 XX  
 XX 08-OCT-2001; 2001GB-00024145.  
 XX

PA (FARB ) BAYER AG.  
 XX  
 XX Munnes M, Gehrman M, Wick M, Schmitz G;  
 PI  
 XX WPI; 2003-403108/38.  
 DR  
 DR N-PSDB; ACA89944.

XX Predicting, diagnosing or prognosing a cardiovascular disease, e.g.  
 PT angina, ischemia, myocardial infarction or atherosclerosis by detection  
 PT of a polynucleotide in a biological sample comprises detecting a  
 PT hybridization complex.  
 XX

XX Claim 3; Page 433-435; 454pp; English.  
 XX  
 CC The invention describes a method of predicting, diagnosing or prognosing  
 CC a cardiovascular disease by detection of a polynucleotide in a biological  
 CC sample comprises hybridising at least one of the polynucleotide to a  
 CC nucleic acid material of a biological sample, thus forming a  
 CC hybridisation complex, and detecting the hybridisation complex. The  
 CC polynucleotides, polypeptides, antisense molecule, antibody and reagent  
 CC are useful for preparing compositions for preventing, predicting or  
 CC diagnosing, or a medicament for treating a cardiovascular disease, e.g.

CC atherosclerosis, ischaemia, angina pectoris, or myocardial infarction.  
 CC This sequence represents a protein identified in the invention a being  
 CC differentially expressed in individuals with cardiovascular disease  
 XX  
 SQ Sequence 482 AA;

Query Match 100.0%; Score 95; DB 6; Length 482;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 |||||  
 DB 139 RIAGRGKLVYDSARHH 156

RESULT 16  
 AAB62100  
 ID AAB62100 standard; protein; 564 AA.

XX AAB62100;  
 AC  
 XX 29-MAY-2001 (first entry)  
 DT  
 XX Human bridging integrator-2 (Bin2) protein.

XX Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
 KW hyperplastic disease; cytostatic; cell growth regulator;  
 KW chromosome 4q22.1.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 23..35  
 FT /note= "specifically claimed fragment"  
 FT Peptide 138..155  
 FT /note= "specifically claimed fragment"  
 XX

XX WO200116158-A2.  
 PN  
 XX 08-MAR-2001.  
 PD  
 XX 30-AUG-2000; 2000WO-US023723.  
 XX  
 XX 31-AUG-1999; 99US-0151554P.  
 XX  
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Prendergast GC, Ge K;  
 PI  
 XX WPI; 2001-235087/24.  
 DR  
 DR N-PSDB; AAF57268.

XX New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
 PT useful for regulating cell growth, and for diagnosing or treating  
 PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
 PT or hepatocarcinoma.

XX Claim 1; Fig 1A-C; 62pp; English.

XX This represents a human bridging integrator-2 (Bin2) protein. The Bin2  
 CC protein can be expressed by standard recombinant methodology. The Bin2  
 CC proteins or peptides are useful in regulating cell growth, cell survival,  
 CC differentiation, endocytosis and actin organization. These peptides or  
 CC proteins are also useful for diagnosing or treating conditions associated  
 CC with inappropriate expression of Bin2, e.g. liver cancer, breast cancer,  
 CC hepatocarcinoma, myeloid and lymphoid leukemias or blood disorders. The  
 CC peptides, nucleic acid sequences or anti-Bin2 antibodies are useful for  
 CC diagnosing inappropriate expression of Bin2. Bin2 is also useful for  
 CC treating disorders associated with excessive Bin1 levels, e.g. liver,  
 CC colorectal, prostate or breast cancers, epithelia cell cancers, melanoma,  
 CC or hyperplastic disease states  
 XX

XX Sequence 564 AA;  
 SQ

Query Match 100.0%; Score 95; DB 4; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
 |||||  
 DB 138 RIAKGRKLVYDSARHH 155

RESULT 17  
 ID ABG21411 standard; protein; 572 AA.  
 XX  
 AC ABG21411;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #21402.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 XX  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS85598.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51770; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridization probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 572 AA;  
 Query Match 100.0%; Score 95; DB 4; Length 572;

Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
 |||||  
 DB 145 RIAKGRKLVYDSARHH 162

RESULT 18  
 ADD45951  
 ID ADD45951 standard; protein; 588 AA.  
 XX  
 AC ADD45951;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein O08839, SEQ ID NO 11623.  
 XX  
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 OS Rattus norvegicus.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 XX  
 PR 01-NOV-2001; 2001US-0346382P.  
 XX  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GEHO) GEN HOSPITAL CORP.  
 XX  
 PA (FARB) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 XX WPI; 2003-268312/26.  
 DR GENBANK; O08839.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PT  
 PS Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC the sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 588 AA;  
Query Match 100.0%; Score 95; DB 7; Length 588;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 1 RIAGRGRKLVYDSARHH 18  
139 RIAGRGRKLVYDSARHH 156  
Db  
RESULT 19  
ADE60848  
ID ADE60848 standard; protein; 588 AA.  
XX  
AC ADE60848;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein O08839, SEQ ID NO 6761.  
XX  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX  
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GHEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX GENBANK; O08839.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that

CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
SQ Sequence 588 AA;  
Query Match 100.0%; Score 95; DB 7; Length 588;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 1 RIAGRGRKLVYDSARHH 18  
139 RIAGRGRKLVYDSARHH 156  
Db  
RESULT 20  
ADE60840  
ID ADE60840 standard; protein; 588 AA.  
XX  
AC ADE60840;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein O08839, SEQ ID NO 6753.  
XX  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX  
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GHEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX GENBANK; O08839.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that



CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 100.0%; Score 95; DB 7; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRKLVDYDSARHH 18  
 |||||  
 Db 139 RIAGRKLVDYDSARHH 156

# RESULT 21

ADE60844  
 ID ADE60844 standard; protein; 588 AA.

XX AC ADE60844;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein O08839, SEQ ID NO 6757.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX FA (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; O08839.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 100.0%; Score 95; DB 7; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRKLVDYDSARHH 18  
 |||||  
 Db 139 RIAGRKLVDYDSARHH 156

# RESULT 22

ADE60852

ID ADE60852 standard; protein; 588 AA.

XX AC ADE60852;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein O08839, SEQ ID NO 6765.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX FA (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; O08839.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 588 AA;  
 CC Query Match 100.0%; Score 95; DB 7; Length 588;  
 CC Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 CC Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
 DB 139 RIAKGRKLVYDSARHH 156  
 |||||

RESULT 23  
 ID ADE60842  
 ID ADE60842 standard; protein; 593 AA.

XX AC ADE60842;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human Protein Q99688, SEQ ID NO 6755.

XX KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX PN W02003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (PAB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; Q99688.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 593 AA;

Query Match 100.0%; Score 95; DB 7; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
 DB 139 RIAKGRKLVYDSARHH 156  
 |||||

RESULT 24

ADD45953

ID ADD45953 standard; protein; 593 AA.

XX AC ADD45953;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein Q99688, SEQ ID NO 11625.

XX KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX PN W02003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (PAB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; Q99688.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 593 AA;  
 Query Match 100.0%; Score 95; DB 7; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 |||||  
 Db 139 RIAGRGKLVYDSARHH 156

RESULT 25  
 ADE60850  
 ID ADE60850 standard; protein; 593 AA.  
 XX AC ADE60850;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human Protein Q99688, SEQ ID NO 6763.  
 XX KW Human; pain; neuronal tissue; gene therapy;  
 XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 XX KW spared nerve injury; SNI; Chung.  
 XX OS Homo sapiens.  
 XX PN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX PF 14-AUG-2002; 2002WO-US025765.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 XX PR 01-NOV-2001; 2001US-0346382P.  
 XX PR 26-NOV-2001; 2001US-0333347P.  
 XX PA (GEO) GEN HOSPITAL CORP.  
 XX PA (FARB) BAYER AG.  
 XX PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 XX DR GENBANK; Q99688.  
 XX PI New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI) in an animal (e.g. gene  
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 593 AA;

Query Match 100.0%; Score 95; DB 7; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 |||||  
 Db 139 RIAGRGKLVYDSARHH 156

RESULT 26  
 ADE60846  
 ID ADE60846 standard; protein; 593 AA.  
 XX AC ADE60846;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human Protein Q99688, SEQ ID NO 6759.  
 XX KW Human; pain; neuronal tissue; gene therapy;  
 XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 XX KW spared nerve injury; SNI; Chung.  
 XX OS Homo sapiens.  
 XX PN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX PF 14-AUG-2002; 2002WO-US025765.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 XX PR 01-NOV-2001; 2001US-0346382P.  
 XX PR 26-NOV-2001; 2001US-0333347P.  
 XX PA (GEO) GEN HOSPITAL CORP.  
 XX PA (FARB) BAYER AG.  
 XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.  
 DR GENBANK; Q99688.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 XX Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 593 AA;  
 SQ  
 Query Match 100.0%; Score 95; DB 7; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RIAGKGRKLVYDSARHH 18  
 Db 139 RIAGKGRKLVYDSARHH 156  
 RESULT 28  
 ADE60854  
 ID ADE60854 standard; protein; 593 AA.  
 XX ADE60854;  
 XX  
 XX 29-JAN-2004 (first entry)  
 DT Human Protein Q99688, SEQ ID NO 6767.  
 DE  
 XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 OS  
 XX WO2003016475-A2.  
 PN  
 XX 27-FEB-2003.  
 PD  
 XX 14-AUG-2002; 2002WO-US025765.  
 PF  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 PR

XX (GEHO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 FI WPI; 2003-268312/26.  
 XX GENBANK; Q99688.  
 DR  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 XX Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 593 AA;  
 SQ  
 Query Match 100.0%; Score 95; DB 7; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RIAGKGRKLVYDSARHH 18  
 Db 139 RIAGKGRKLVYDSARHH 156  
 RESULT 28  
 ABG20887  
 ID ABG20887 standard; protein; 594 AA.  
 XX ABG20887;  
 XX  
 XX 13-FEB-2002 (first entry)  
 DT Novel human diagnostic protein #20878.  
 DE  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW Homo sapiens.  
 OS  
 XX WO200175067-A2.  
 PN  
 XX 11-OCT-2001.  
 PD  
 XX 30-MAR-2001; 2001WO-US008631.  
 PF

XX 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.  
XX N-PSDB; AAS85074.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 20; SEQ ID NO 51246; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 594 AA;

Query Match 100.0%; Score 95; DB 4; Length 594;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RIAGRGKLVYDSARHH 18  
DB 167 RIAGRGKLVYDSARHH 184

RESULT 29

ABG20324

ID ABG20324 standard; protein; 594 AA.

XX ABG20324;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #20315.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS84511.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 20; SEQ ID NO 50893; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 594 AA;

Query Match 100.0%; Score 95; DB 4; Length 594;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18

DB 167 RIAGRGKLVYDSARHH 184

RESULT 30

ABG22457

ID ABG22457 standard; protein; 594 AA.

XX ABG22457;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #22448.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS86644.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 52816; 103pp; English.  
 XX  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 594 AA;  
 SQ  
 Query Match 100.0%; Score 95; DB 4; Length 594;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RIAKRGKLVYDSARHH 18  
 DB 167 RIAKRGKLVYDSARHH 184  
 RESULT 31  
 ABG20886  
 ID ABG20886 standard; protein; 960 AA.  
 XX  
 AC ABG20886;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #20877.  
 XX  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW  
 OS Homo sapiens.  
 XX  
 OS WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PD 30-MAR-2001; 2001WO-US008631.  
 XX  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX

PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS85073.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 51245; 103pp; English.  
 XX  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 960 AA;  
 SQ  
 Query Match 100.0%; Score 95; DB 4; Length 960;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RIAKRGKLVYDSARHH 18  
 DB 238 RIAKRGKLVYDSARHH 255  
 RESULT 32  
 ABG20323  
 ID ABG20323 standard; protein; 960 AA.  
 XX  
 AC ABG20323;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #20314.  
 XX  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW  
 OS Homo sapiens.  
 XX  
 OS WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PD 30-MAR-2001; 2001WO-US008631.  
 XX  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX

DR WPI: 2001-639362/73.  
DR N-PSDB; AAS84510.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 20; SEQ ID NO 50682; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activities. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 960 AA;  
  
Query Match 100.0%; Score 95; DB 4; Length 960;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RIAGRGKLVYDSARHH 18  
Db 238 RIAGRGKLVYDSARHH 255  
  
RESULT 33  
ADD44887  
ID ADD44887 standard; protein; 683 AA.  
AC ADD44887;  
XX  
DT 29-JAN-2004 (first entry)  
DE Rat Protein O08838, SEQ ID NO 10318.  
XX  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-033347P.  
XX  
XX (GEO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX

DR WPI: 2003-268312/26.  
DR GENBANK; O08838.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 683 AA;  
  
Query Match 93.7%; Score 89; DB 7; Length 683;  
Best Local Similarity 94.4%; Pred. No. 1.7e-06;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RIAGRGKLVYDSARHH 18  
Db 134 RIAGRGKLVYDSARHH 151  
  
RESULT 34  
AAE22091  
ID AAE22091 standard; protein; 695 AA.  
XX  
AC AAE22091;  
XX  
DT 25-JUL-2002 (first entry)  
DE Human amphiphsin-1 protein.  
XX  
DE Cellular senescence; amphiphsin-1 protein; caveolin-1 protein; human;  
XX gene therapy.  
KW Homo sapiens.  
XX  
XX WO200221140-A1.  
XX  
XX 14-MAR-2002.  
XX  
XX 06-JUL-2001; 2001WO-KR001159.  
XX  
XX 08-SEP-2000; 2000KR-00053341.  
XX 08-SEP-2000; 2000KR-00053342.  
XX  
XX (META-) METABOLIC ENG LAB CO LTD.  
XX

PI Park S, Park W, Park J, Cho K, Kim D;  
 XX WPI; 2002-362263/39.  
 DR N-PSDB; RAD35148.  
 XX  
 PT Modulating cellular senescence in patient involves administering protein  
 PT involved in cellular senescence e.g., amphiphysin protein or caveolin  
 PT protein or polynucleotide encoding the proteins.  
 XX  
 PS Claim 23; Page 92-95; 103pp; English.  
 XX  
 CC The invention relates to a method of modulating cellular senescence in a  
 CC patient. The method involves administering protein involved in cellular  
 CC senescence e.g. amphiphysin-1 protein or caveolin-1 protein or  
 CC polynucleotides encoding such proteins. The method is useful for  
 CC detecting and modulating cellular senescence in a mammalian cell. It is  
 CC also used in gene therapy. The present sequence is human amphiphysin-1  
 CC protein  
 XX Sequence 695 AA;  
 SQ  
 Query Match 93.7%; Score 89; DB 5; Length 695;  
 Best Local Similarity 94.4%; Pred. No. 1.8e-06;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RIAKGRKLVYDSARHH 18  
 Db 134 RIAKGRKLVYDSARHH 151  
 RESULT 35  
 AAU75110  
 ID AAU75110 standard; protein; 695 AA.  
 AC AAU75110;  
 XX  
 DT 09-APR-2002 (first entry)  
 DE  
 XX Synaptic vesicle and endocytosis associated protein, amphiphysin.  
 XX  
 KW MAPKAP-K3; AP-3 delta; APP-695; amyloid A4; Hsp8; heat shock protein 8;  
 KW L130; NY-REN-58; P38 Alpha; P38 Beta; ERK3; KIAA0934; CDK9;  
 KW cell cycle dependent kinase 9; CLK; C-NAP-1; clathrin heavy chain;  
 KW amphiphysin; PN9109; KIAA1106; neurodegenerative disorder;  
 KW non-insulin dependent diabetes mellitus; NIDDM; Alzheimer's disease;  
 KW inflammatory disease; rheumatoid arthritis; inflammatory bowel disorder;  
 KW atherosclerosis; cardiac hypertrophy; hypoxic brain injury;  
 KW yeast two-hybrid; signal transduction pathway; human; endocytosis;  
 KW synaptic vesicle; mitogen activated protein kinase.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 93..273  
 FT /note="This region binds the centrosomal Nek-2  
 FT associated protein 1 (C-NAP1) (see ABK13313), the bait  
 FT protein in a yeast two-hybrid assay, producing a complex  
 FT Claimed in claim 1"  
 XX  
 PN WO200198524-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 XX 21-JUN-2001; 2001WO-US019762.  
 PF  
 XX 22-JUN-2000; 2000US-0213245P.  
 PR  
 XX (MYRI-) MYRIAD GENETICS INC.  
 PA Heichman K, Bartel PL;  
 PI WPI; 2002-122287/16.  
 XX  
 XX WPI; 2002-122287/16.

PT New protein complexes comprising protein-protein interactions (e.g.  
 PT MAPKAP-K3/AP-3 delta or C-NAP-1/clathrin HC), useful for diagnosing  
 PT physiological generative disorders or screening drugs for these diseases.  
 XX  
 XX Example 11; Page: 60pp; English.  
 PS  
 CC The invention describes an isolated protein complex, comprising two  
 CC proteins. The protein complex comprises: protein kinase MAPKAP-K3 and AP-  
 CC 3 adaptor complex delta sub-unit; MAPKAP-K3 and amyloid A4 precursor  
 CC protein, APP-695; MAPKAP-K3 and heat shock protein (Hsp) 8; leucine rich  
 CC L130 and NY-REN-58; P38 Alpha and P38 Beta; protein kinase ERK3 and  
 CC KIAA0934 (unknown function); ERK3 and cell cycle dependent kinase (CDK) 9  
 CC ; ERK3 and protein kinase CLK; C-NAP-1 and Clathrin heavy chain; C-NAP-1  
 CC and Amphiphysin; C-NAP-1 and novel protein PN9109 or C-NAP-1 and KIAA1106  
 CC (unknown function) interactions. The protein complexes are useful for  
 CC diagnosing physiological generative disorders, drug screening for agents  
 CC that modulate the interaction of the proteins (thus identify drug  
 CC targets), and identifying additional proteins in the pathway common to  
 CC the proteins. These physiological disorders include non-insulin dependent  
 CC diabetes mellitus (NIDDM), neurodegenerative disorders (e.g. Alzheimer's  
 CC disease), inflammatory diseases (e.g. rheumatoid arthritis and  
 CC inflammatory bowel disorder) and other human disease such as  
 CC atherosclerosis, cardiac hypertrophy and hypoxic brain injury. This  
 CC sequence represents the synaptic vesicle and endocytosis associated  
 CC protein amphiphysin, residues 93-273 of which binds to the bait protein  
 CC centrosomal Nek-2 associated protein 1 (C-NAP1) (see ABK13313) in a yeast  
 CC two-hybrid assay for determining components of signal transduction  
 CC pathways and forms an interaction claimed in claim 1 of the invention.  
 CC Note: This sequence does not appear in the specification but has been  
 CC obtained from a reference given in the invention  
 XX  
 XX Sequence 695 AA;  
 SQ  
 Query Match 93.7%; Score 89; DB 5; Length 695;  
 Best Local Similarity 94.4%; Pred. No. 1.8e-06;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RIAKGRKLVYDSARHH 18  
 Db 134 RIAKGRKLVYDSARHH 151  
 RESULT 36  
 ADD44889  
 ID ADD44889 standard; protein; 695 AA.  
 XX ADD44889;  
 XX  
 DT 29-JAN-2004 (first entry)  
 DE  
 XX Human Protein P49418, SEQ ID NO 10320.  
 XX  
 XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 XX 14-AUG-2002; 2002WO-US025765.  
 PF  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR  
 XX 01-NOV-2001; 2001US-0346382P.  
 PR  
 XX 26-NOV-2001; 2001US-0333347P.  
 PR  
 XX (GEHO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2002-122287/16.





CC plant nucleotide sequence having an open reading frame that encodes a  
CC polypeptide associated with disease resistance or its fragment having  
CC substantially the same activity as the full-length polypeptide. The  
CC polynucleotide of the invention is useful for conferring resistance or  
CC tolerance to a plant pathogen. The present sequence represents a protein  
CC conferring disease resistance used in the invention.

XX Sequence 276 AA;

Query Match 51.6%; Score 49; DB 6; Length 276;  
Best Local Similarity 44.4%; Pred. No. 4.9;  
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RIAGKRGKLVYDSARHH 18  
Db 249 RVRRSTRLTNYDSMERH 266

RESULT 39

ABB91883  
ID ABB91883 standard; protein; 424 AA.

XX ABB91883;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1094.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (PARB ) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences  
XX from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms.

XX Claim 5; SEQ ID NO 1094; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)  
XX for herbicidally active compounds, comprising aligning and comparing  
XX nucleic acid or amino acid sequences from plant with nucleic acid or  
XX amino acid sequences from non-plant organisms using suitable search  
XX parameters, where plant sequences having an E-value greater by a factor  
XX of 3 than the E-value of most similar non-plant sequences are selected.  
XX The polypeptides or nucleic acids encoding them are useful for  
XX identifying modulators. The identified modulators are useful as  
XX herbicides

XX Sequence 424 AA;

Query Match 49.5%; Score 47; DB 5; Length 424;  
Best Local Similarity 53.3%; Pred. No. 17;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARHH 18

Db 382 KEGKKFVDINKFRHH 396

RESULT 40

AAM48428  
ID AAM48428 standard; protein; 425 AA.

XX AAM48428;

XX 16-JUL-2002 (first entry)

XX Protein ARSKL.

XX Phosphorylase AtsIK; osmotic stress inducing gene.

XX Unidentified.

XX KR2001035308-A.

XX 07-MAY-2001.

XX 02-FEB-2001; 2001KR-00005097.

XX 02-FEB-2001; 2001KR-00005097.

XX (GENO-) GENOMINE INC.

XX Hwang IH, Lim JH, Pi GT;

XX WPI; 2001-594308/67.

XX Phosphorylase atsik inhibiting expression of osmotic stress inducing  
XX gene.

XX Disclosure; Fig 1; 20pp; Korean.

XX The present invention relates to phosphorylase AtsIK (AAM48428), which  
XX inhibits an osmotic stress inducing gene in order to further increase the  
XX productivity of plants with greater resistance to the stress.  
XX Phosphorylase AtsIK is obtained by: i) cultivating Arabidopsis thaliana  
XX on MS plate at 22 deg.C in culture room or in greenhouse where the sample  
XX is placed under bright condition for 16 hours and under dark condition  
XX for 8 hours, repeatedly, with 70% moisture; ii) manufacturing a  
XX subtraction library by comparing cDNA extracted from plant induced by  
XX salt stress with control cDNA and eliminating cDNA being always expressed  
XX ; iii) collecting cDNA from the library randomly for base sequencing and  
XX obtaining 4 cDNA segments; and iv) selecting cDNA clones whose expression  
XX is controlled by osmotic stress and, selecting again a 2090bp cDNA which  
XX has an open reading frame encoding 557 amino acids, among the clones. The  
XX present protein sequence was used in a sequence homology alignment with  
XX the sequence of AtsIK

XX Sequence 425 AA;

Query Match 49.5%; Score 47; DB 4; Length 425;  
Best Local Similarity 53.3%; Pred. No. 17;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARHH 18

Db 383 KEGKKFVDINKFRHH 397

RESULT 41

AAG41757  
ID AAG41757 standard; protein; 468 AA.

XX AAG41757;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 51993.

XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.  
XX Arabidopsis thaliana.  
XX EPI033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 26-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 28-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0132048P.  
XX 04-MAY-1999; 99US-0132407P.  
XX 05-MAY-1999; 99US-0132484P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136332P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137282P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139492P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
XX 18-JUN-1999; 99US-0139462P.  
XX 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
XX 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
XX 23-JUN-1999; 99US-0140354P.  
XX 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151308P.  
PR 01-SEP-1999; 99US-0151330P.  
PR 07-SEP-1999; 99US-0152363P.

```

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 28-OCT-1999; 99US-0162142P.

Query Match 49.5%; Score 47; DB 3; Length 468;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 4 KRGRKLVYDSAEHH 18
Db 166 KAGREVDYLGTRHH 180

RESULT 42
AAG41756
ID AAG41756 standard; protein; 472 AA.
XX
AC AAG41756;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51992.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123150P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139464P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140821P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 05-JUL-1999; 99US-0142390P.

```

PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142820P.  
 PR 12-JUL-1999; 99US-0142377P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 17-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.

PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160988P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161366P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 49.5%; Score 47; DB 3; Length 472;  
 Best Local Similarity 53.3%; Pred. No. 20;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARHH 18  
 |||:|  
 Db 170 KAGREVDYLGTRHH 184

RESULT 43  
 AAG41755  
 ID AAG41755 standard; protein; 587 AA.

XX AC AAG41755;  
 XX DT 18-OCT-2000 (first entry)  
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 51991.  
 XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX KW termination sequence.

XX OS Arabidopsis thaliana.  
 XX FN EP1033405-A2.  
 XX PD 06-SEP-2000.

XX EF 25-FEB-2000; 2000EP-00301439.  
 XX PR 25-FEB-1999; 99US-0121825P.  
 XX PR 05-MAR-1999; 99US-0123180P.  
 XX PR 09-MAR-1999; 99US-0123548P.  
 XX PR 23-MAR-1999; 99US-0125788P.

PR	25-MAR-1999;	99US-0126264P.	PR	16-JUL-1999;	99US-0144086P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144325P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144331P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144332P.
PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	21-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132040P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145132P.
PR	07-MAY-1999;	99US-0132487P.	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134218P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134219P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134222P.	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	99US-0134768P.	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	99US-0134941P.	PR	02-AUG-1999;	99US-0146386P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146389P.
PR	24-MAY-1999;	99US-0135629P.	PR	03-AUG-1999;	99US-0147038P.
PR	25-MAY-1999;	99US-0136021P.	PR	04-AUG-1999;	99US-0147204P.
PR	27-MAY-1999;	99US-0136392P.	PR	05-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	99US-0136782P.	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	99US-0137222P.	PR	06-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147303P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147416P.
PR	07-JUN-1999;	99US-0137724P.	PR	09-AUG-1999;	99US-0147493P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138540P.	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	99US-0138847P.	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	99US-0139119P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148565P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	99US-0139492P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139454P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139455P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139456P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139459P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139461P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139462P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139463P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140695P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0140991P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155565P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142977P.	PR	05-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143424P.	PR	08-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0143624P.	PR	07-OCT-1999;	99US-0158029P.
PR	15-JUL-1999;	99US-0144005P.			
PR	16-JUL-1999;	99US-0144085P.			

PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 49.5%; Score 47; DB 3; Length 587;  
 Best Local Similarity 53.3%; Pred. NO. 25;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KGRKLVYDSARHH 18  
 |||||  
 DB 285 KAGREVADVLGTRHH 299

RESULT 44  
 ABU24499  
 ID ABU24499 standard; protein; 765 AA.  
 XX AC ABU24499;  
 XX DT 19-JUN-2003 (first entry)  
 XX DE Protein encoded by Prokaryotic essential gene #10026.  
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX OS Clostridium botulinum.  
 XX PN WO20027183-A2.  
 XX PD 03-OCT-2002.  
 XX PF 21-MAR-2002; 2002WO-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen K, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA28369.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 52423; 1766pp; English.  
 XX

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *X. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 765 AA;

Query Match 49.5%; Score 47; DB 6; Length 765;  
 Best Local Similarity 61.5%; Pred. NO. 33;  
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KGRKLVYDSAR 16  
 |||||  
 DB 422 KGRKLVYDSAR 434

RESULT 45  
 ABP55425  
 ID ABP55425 standard; protein; 521 AA.  
 XX AC ABP55425;  
 XX DT 04-FEB-2003 (first entry)  
 XX DE Mouse polycomb gene enhancer 84-57.31 protein SEQ ID NO:2.  
 XX KW Mouse; polycomb gene enhancer 84-57.31; embryonic development deformity;  
 XX OS Mus sp.  
 XX PN CN1342699-A.  
 XX PD 03-APR-2002.  
 XX PF 12-SEP-2000; 2000CN-00125168.  
 XX PR 12-SEP-2000; 2000CN-00125168.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX Mao Y, Xie Y;

XX WPI; 2002-529776/57.

DR N-PSDB; ABQ83917.

XX A novel mouse polycarb gene enhancer 84-57.31 polypeptide, and the  
PT polynucleotide encoding it, useful for treating several diseases e.g.  
PT embryonic development deformity and tumors.

XX Claim 1; Page 26-27 (Disclosure); 34pp; Chinese.

XX The present sequence represents mouse polycarb gene enhancer 84-57.31  
CC (1). Also described is a process for preparing (1) using DNA  
CC recombination techniques. (1) can be used for treating several diseases  
CC e.g. embryonic development deformity and tumors

XX Sequence 521 AA;

Query Match 47.9%; Score 45.5; DB 5; Length 521;

Best Local Similarity 43.5%; Pred. No. 40;

Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

QY 1 RIAKRGKLV-----DYDSARHH 18

Db 143 RVGRGVRLLDRAHSDYDSVPHH 165

RESULT 46

ABU96724

ID ABU96724 standard; protein; 578 AA.

XX AC ABU96724;

XX 25-JUL-2003 (first entry)

XX Human nucleic acid-associated protein (NAAP) #33.

XX Human; nucleic acid-associated protein; cytostatic; antiarteriosclerotic;  
KW anticonvulsant; nontropic; neuroprotective; cerebroprotective; anti-HIV;  
KW anti-allergic; anti-inflammatory; thyromimetic; gene therapy;  
KW cell proliferative disorder; cancer; atherosclerosis;  
KW neurological disorder; epilepsy; Huntington's disease; stroke;  
KW immune disorder; inflammatory disorder; AIDS; allergy;  
KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;  
KW protein-protein interaction; drug-target interaction;  
KW gene expression profile.

XX Homo sapiens.

XX WO2003023003-A2.

XX 20-MAR-2003.

XX 05-SEP-2002; 2002WO-US028540.

XX 07-SEP-2001; 2001US-0317792P.

XX 07-SEP-2001; 2001US-0317912P.

XX 14-SEP-2001; 2001US-032270P.

XX 21-SEP-2001; 2001US-0324040P.

XX 28-SEP-2001; 2001US-0326732P.

XX 19-OCT-2001; 2001US-0346716P.

XX 25-JAN-2002; 2002US-0351749P.

XX 22-FEB-2002; 2002US-0359498P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Jackson JL, Griffin JA, Elliott VS, Forsythe IJ;

XX Becha SD, Richardson TW, Lee EA, Sprague WW, Emerling BM;

XX Thangavelu K, Warren BA, Tran UK, Yue H, Xu Y, Yue H, Li JX;

XX Hafalia AJA, Sanjanwala B, Marquis JP, Corvad AE, Lee SX, Ison CH; P;

XX Baughn MR, Chawla NK, Nguyen DB, Swarnakar A, Zebarjadian Y, Shah P;

PI

XX

PI

XX

DR

XX

XX

PT

PT

PT

PT

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

XX

SQ

Sequence 578 AA;

Query Match 47.9%; Score 45.5; DB 6; Length 578;

Best Local Similarity 43.5%; Pred. No. 44;

Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

QY 1 RIAKRGKLV-----DYDSARHH 18

Db 408 RVGRGVRLLDRAHSDYDSVPHH 430

RESULT 47

AAB95815

ID AAB95815 standard; protein; 596 AA.

XX AAB95815;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:18813.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



DR WPI; 2001-318749/34.  
 XX  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
 PT  
 XX  
 XX  
 PS Claim 8; SEQ ID NO 18813; 2537pp + Sequence Listing; English.  
 XX  
 XX  
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer, and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the polynucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention  
 XX  
 XX  
 SQ Sequence 596 AA;  
 Query Match 47.9%; Score 45.5; DB 4; Length 596;  
 Best Local Similarity 43.5%; Pred. No. 46;  
 Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;  
 QY 1 RIAGRGKLV-----DYDSARHH 18  
 DB 241 RVGRGRLLDRAHSDYDSVFHH 263  
 RESULT 48  
 ADB64808  
 ID ADB64808 standard; protein; 742 AA.  
 XX  
 AC ADB64808;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Human protein encoded by clone OCBEP20016810.  
 XX  
 KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration; cell regeneration; membrane protein; signal transduction-related protein; transcription-related protein; osteoporosis; neurological disease; cancer; tumour.  
 KW  
 XX Homo sapiens.  
 OS  
 XX EPI308459-A2.  
 FN  
 XX 07-MAY-2003.  
 PD  
 XX 28-MAR-2002; 2002EP-00007401.  
 PF  
 XX 05-NOV-2001; 2001JP-00379298.  
 PR  
 XX 25-JAN-2002; 2002US-00350978.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Sugiyama T, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR WPI; 2003-450961/43.  
 DR N-PSDB; AD362838.  
 XX  
 PT New polynucleotides and polypeptides, useful for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy.  
 PT  
 XX  
 PS Claim 1; Page; 222pp; English.  
 XX  
 XX  
 CC The invention discloses a polynucleotide comprising a sequence selected from 1970 fully defined nucleotide sequences which encode novel polypeptides. Also claimed is a polypeptide encoded by the polynucleotide or its partial peptide, an antibody binding to the polypeptide or peptide of the polynucleotide, immunologically assaying the polypeptide or peptide of the polynucleotide by contacting the polypeptide or peptide with the antibody of the encoded protein, and observing the binding between the two, a transformant carrying the polynucleotide in an expressible manner and an antisense polynucleotide. The oligonucleotide is useful as a primer for synthesising the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a protein of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the CC European Patent Office.  
 CC  
 XX  
 SQ Sequence 742 AA;  
 Query Match 47.9%; Score 45.5; DB 7; Length 742;  
 Best Local Similarity 43.5%; Pred. No. 59;  
 Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;  
 QY 1 RIAGRGKLV-----DYDSARHH 18  
 DB 387 RVGRGRLLDRAHSDYDSVFHH 409  
 RESULT 49  
 ABU00410  
 ID ABU00410 standard; protein; 763 AA.  
 XX  
 AC ABU00410;  
 XX  
 DT 17-JAN-2003 (first entry)  
 XX  
 DE Human novel polypeptide #503.  
 XX  
 KW Human; genetic disorder; gene mapping; medical imaging; cancer; neurodegenerative disorder; lymphoid cell disorder; osteoporosis; Parkinson's disease; Alzheimer's disease; bone degenerative disorder; osteoarthritis; periodontal disease; liver fibrosis; viral infection; fungal infection; bacterial infection; autoimmune disease; diabetes; atopic dermatitis.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200274961-A1.  
 FN  
 XX 26-SEP-2002.  
 PD  
 XX 14-MAR-2002; 2002WO-US005109.  
 PF  
 XX

PR 15-MAR-2001; 2001US-00810173.  
 PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
 PT Wehrman T, Wang J, Wang D, Drmanac RT;  
 DR WPI; 2003-040556/03.  
 XX N-PSDB; ABX05488.  
 PT New isolated polypeptides and polynucleotides, useful for preventing,  
 PT treating or ameliorating medical conditions, such as cancer,  
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
 PT disorders, and infections.  
 PS Claim 9; SEQ ID NO 1029; 235pp; English.  
 XX The invention relates to human polynucleotides and the polypeptides they  
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
 CC forensics, gene mapping, medical imaging, identification of mutations,  
 CC and producing many other types of data and traits, assessing biodiversity,  
 CC amino acid sequences. They are also useful for products dependent on DNA and  
 CC ameliorating medical conditions. They are also useful for preventing, treating or  
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease, lymphoid cell  
 CC pericardial disease, osteoporosis, osteoarthritis, bone degenerative  
 CC bacterial disease, liver fibrosis, infections (e.g. viral, fungal or  
 CC Sequences ABG9988-RBG9989 and ABU0010-ABU00433 represent human  
 CC polypeptides of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied by the European Patent Office  
 XX Sequence 763 AA;  
 SQ

Query Match  
 Best Local Similarity 47.9%; Score 45.5; DB 6; Length 763;  
 Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;  
 QY 1 RIAKGRKLV-----DVDSARHH 18  
 Db 408 RVGGRGVLLDRAHSYDSVFFH 430  
 RESULT 50  
 AAU63157  
 ID AAU63157 standard; protein; 161 AA.  
 AC AAU63157;  
 XX  
 DT 27-FEB-2002 (first entry)  
 DE Propionibacterium acnes immunogenic protein #24053.  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 OS Propionibacterium acnes.  
 XX WO200181581-A2.  
 PN  
 PD 01-NOV-2001.  
 XX  
 PP 20-APR-2001; 2001WO-US012865.  
 XX  
 PR 21-APR-2000; 2000US-0199047P.  
 PR 02-JUN-2000; 2000US-0208841P.  
 PR 07-JUL-2000; 2000US-0216747P.  
 XX  
 PA (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX WPI; 2001-616774/71.  
 DR N-PSDB; AAS59632.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.  
 PS Example 1; SEQ ID NO 24352; 1069pp; English.  
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis,  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies may also be used as  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may be used to  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 161 AA;  
 SQ

Query Match  
 Best Local Similarity 46.3%; Score 44; DB 4; Length 161;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 6 GRKLVYDSARHH 18  
 Db 27 GGELIALDSARHH 39

Search completed: March 4, 2004, 17:44:23  
 Job time : 64.6452 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:39:06 ; Search time 15.6774 Seconds

(without alignments)  
110.442 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_138\_155

Perfect score: 95

Sequence: 1 RIAGRGKLVVDVSARRH 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : PIR 78.\*

1: piri.\*

2: piri2.\*

3: piri3.\*

4: piri4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	482	JC5593	amphiphysin II2 -
2	89	93.7	682	S22700	amphiphysin - chic
3	89	93.7	695	S62400	amphiphysin (clone
4	69	72.6	461	T22946	hypothetical prote
5	49.5	52.1	130	F91120	hypothetical prote
6	48	50.5	159	F84259	hypothetical prote
7	47	49.5	424	F84658	probable protein k
8	47	49.5	578	T51888	asparagine synthas
9	47	49.5	578	T50812	asparagine synthas
10	45	47.4	288	S40975	hypothetical prote
11	44	46.3	225	PC2237	protein-glutamine
12	43.5	45.8	282	AB0828	probable transcrip
13	43	45.3	284	T40661	yeast reduced viab
14	43	45.3	270	T43000	RV5161 protein hom
15	43	45.3	397	S62579	probable plasma me
16	43	45.3	848	I55498	testicular dynamin
17	42	44.2	252	B69497	conserved hypotet
18	42	44.2	463	D70933	hypothetical prote
19	42	44.2	906	S32607	trifunctional enzy
20	42	44.2	1133	F83264	hypothetical prote
21	42	44.2	1957	A59294	skeletal myosin -
22	41.5	43.7	731	1 MNVUSH	cucumisn (EC 3.4.
23	41	43.2	92	1	nonstructural prot
24	41	43.2	136	T13595	hypothetical prote
25	41	43.2	237	E70945	hypothetical prote
26	41	43.2	402	T25732	hypothetical prote
27	41	43.2	407	C71055	probable RNA methy
28	41	43.2	974	D89057	protein K09H11.1
29	41	43.2	1001	2 T50914	hypothetical membr

30	40.5	42.6	306	1	H65033	hypothetical prote
31	40.5	42.6	306	2	C91057	hypothetical prote
32	40.5	42.6	306	2	G85901	hypothetical prote
33	40.5	42.6	315	2	A71288	probable methylene
34	40.5	42.6	393	2	AC2896	oxidoreductase Atu
35	40.5	42.6	408	2	F97671	probable oxidoredu
36	40.5	42.6	458	2	S77016	sensory transducti
37	40	42.1	225	2	B35387	hypothetical prote
38	40	42.1	264	2	AG3175	hypothetical prote
39	40	42.1	308	2	E81288	hypothetical prote
40	40	42.1	426	2	H70390	conserved hypotet
41	40	42.1	651	2	T49986	lectin-like protei
42	40	42.1	752	2	S56146	GCN20 protein - ye
43	40	42.1	830	2	T18860	hypothetical prote
44	39.5	41.6	229	2	H64367	dolichyl-phosphate
45	39	41.1	97	2	T31023	conserved hypotet
46	39	41.1	99	2	G89986	hypothetical prote
47	39	41.1	165	2	A86199	hypothetical prote
48	39	41.1	219	2	T01186	protein kinase (EC
49	39	41.1	224	2	T01185	protein kinase (EC
50	39	41.1	259	2	T49291	hypothetical prote
51	39	41.1	394	2	AF2650	conserved hypotet
52	39	41.1	420	2	E97432	hypothetical prote
53	39	41.1	459	2	E71443	probable DNA-bindi
54	39	41.1	461	2	B95388	Probable (EC 1.1.1
55	39	41.1	512	2	S48828	lethal-3 protein,
56	39	41.1	559	2	T02825	probable membrane
57	39	41.1	639	2	B95945	probable glycosyl
58	39	41.1	756	2	T05829	hypothetical prote
59	39	41.1	781	2	T05206	hypothetical prote
60	39	41.1	796	2	C85220	hypothetical prote
61	39	41.1	918	2	S04255	regulatory protein
62	39	41.1	980	2	A38523	genome polyprotein
63	39	41.1	981	2	S55132	hypothetical prote
64	39	41.1	2894	2	C64474	hypothetical prote
65	38.5	40.5	64	2	C87511	hypothetical prote
66	38.5	40.5	427	2	A84155	hypothetical prote
67	38	40.0	136	2	A11104	B. subtilis YacZ p
68	38	40.0	136	2	A11466	B. subtilis YacZ p
69	38	40.0	203	2	JC1247	GRP-binding protei
70	38	40.0	224	2	B81783	hypothetical prote
71	38	40.0	252	2	F87259	hypothetical prote
72	38	40.0	299	2	G82393	transcription regu
73	38	40.0	389	2	D86424	43.3K hypotetical
74	38	40.0	431	2	S45038	protein disulfide-
75	38	40.0	440	2	JC4369	P5 protein precurs
76	38	40.0	464	2	B64970	Colanic acid biosy
77	38	40.0	464	2	H90984	hypothetical prote
78	38	40.0	464	2	C95830	hypothetical prote
79	38	40.0	503	2	B82880	multiple banded an
80	38	40.0	504	2	F71253	conserved hypotet
81	38	40.0	573	2	H99904	aerobic glycerol-3
82	38	40.0	620	2	H85431	Arpase-like protei
83	38	40.0	642	2	S18667	deoxyribodipyrimid
84	38	40.0	651	2	S35706	NADH oxidase (EC 1
85	38	40.0	715	2	A41511	staphylocoagulase
86	38	40.0	757	2	F87304	beta-N-acetylhexos
87	38	40.0	950	2	G83167	valyl-tRNA synthet
88	38	40.0	1433	2	G01946	nitric-oxide synth
89	38	40.0	1518	2	S37928	probable purine nu
90	38	40.0	2429	1	SJHUA	spectrin alpha cha
91	38	40.0	4574	2	G02520	plectin - human
92	38	40.0	4684	2	A59404	plectin [imported]
93	37.5	39.5	347	2	S33939	minor core protein
94	37.5	39.5	366	2	E72355	hypothetical prote
95	37.5	39.5	453	2	B97738	outer membrane pro
96	37.5	39.5	697	2	T00267	hypothetical prote
97	37.5	39.5	1249	2	T36294	hypothetical prote
98	37	38.9	89	1	H64116	ribosomal protein
99	37	38.9	91	2	F82522	hypothetical prote
100	37	38.9	112	2	T34589	hypothetical prote
101	37	38.9	113	2	S43583	F26F3.3 protein -
102	37	38.9	128	2	E69306	hypothetical prote

103 37 38.9 145 1 S23063 ribosomal protein  
104 37 38.9 147 2 E90018 50S ribosomal prot  
105 37 38.9 149 2 T16440 hypothetical prote  
106 37 38.9 150 1 OKYS1 cell division cont  
107 37 38.9 163 2 S28136 gas vesicle protei  
108 37 38.9 200 2 F71866 hypothetical prote  
109 37 38.9 213 2 G82723 partition protein  
110 37 38.9 227 2 AE2343 hypothetical prote  
111 37 38.9 245 2 F83270 probable nucleosid  
112 37 38.9 258 2 T09636 yscE protein - Lac  
113 37 38.9 278 2 C72807 probable non-heme  
114 37 38.9 315 2 S17952 acyltransferase (E  
115 37 38.9 338 2 C75161 3-dehydroquinase  
116 37 38.9 355 2 C64582 phospholipase A1 p  
117 37 38.9 355 2 H71930 probable phospholi  
118 37 38.9 369 2 H88535 protein B0523.4 [i  
119 37 38.9 376 2 JN0745 site-specific DNA-  
120 37 38.9 389 2 E90431 sulfolipid biosynt  
121 37 38.9 392 2 F69544 conserved hypotet  
122 37 38.9 394 2 C82951 conserved hypotet  
123 37 38.9 397 2 G97075 probable transcrip  
124 37 38.9 422 2 S58173 ketoacyl synthase  
125 37 38.9 431 2 S19656 protein disulfide-  
126 37 38.9 439 2 S44733 b0523.4 protein -  
127 37 38.9 539 2 A64441 O-sialoglycoprotei  
128 37 38.9 541 2 S76017 hypothetical prote  
129 37 38.9 543 2 G87598 Tidd/FmbA family p  
130 37 38.9 565 2 T47775 hypothetical prote  
131 37 38.9 586 2 T02978 asparagine synthas  
132 37 38.9 662 1 A49882 histidine decarbox  
133 37 38.9 672 2 T12524 hypothetical prote  
134 37 38.9 797 2 A70453 glutamate ammonia  
135 37 38.9 808 2 F84038 phenylalanyl-tRNA  
136 37 38.9 900 2 A72524 probable alanyl-tr  
137 37 38.9 1139 2 S28277 hypothetical prote  
138 37 38.9 1211 2 T27522 hypothetical prote  
139 37 38.9 1526 2 T19473 hypothetical prote  
140 37 38.9 1670 2 S71551 DNA-directed DNA p  
141 36.5 38.4 124 2 F87429 hypothetical prote  
142 36.5 38.4 439 2 E95858 conserved hypotet  
143 36.5 38.4 431 2 E81053 seryl-tRNA synthet  
144 36.5 38.4 431 2 B81822 serine-tRNA ligase  
145 36.5 38.4 431 2 H84069 hypothetical prote  
146 36.5 38.4 584 2 S06318 endoplasmic reticu  
147 36.5 38.4 638 1 ISMSER protein disulfide-  
148 36.5 38.4 643 1 S32476 protein disulfide-  
149 36 37.9 87 2 T16008 hypothetical prote  
150 36 37.9 89 1 R3EC15 ribosomal protein

# ALIGNMENTS

RESULT 1  
JC5593  
amphiphysin II2 - human  
C:Species: Homo sapiens (man)  
C>Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 09-Jun-2000  
C/Accession: JC5593  
R;Tsuboi, K.; Maeda, Y.; Tsuboi, K.; Seki, S.; Tokunaga, A.  
Biochem. Biophys. Res. Commun. 236, 178-183, 1997  
A;Title: cDNA cloning of a novel amphiphysin isoform and tissue-specific expression of i  
A;Reference number: JC5593; MUID:97366618; PMID:9223448  
A;Accession: JC5593  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-482 <RSU>  
A;Cross-references: DBJ:AF001383; NID:g2199534; PID:AAB61363.1; PID:g2199535  
A;Experimental source: fetal brain  
C;Comment: This protein is involved in the synaptic vesicle recycling and in the regulat  
C;Superfamily: amphiphysin; RVS161 protein homology  
F;16-275/Domain: RVS161 protein homology <RVS>  
F;410-481/Domain: SH3 #status predicted <SH3>

Query Match 100.0%; Score 95; DB 2; Length 482;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RIAXGRKLVYDSARHH 18  
DB 139 RIAXGRKLVYDSARHH 156  
RESULT 2  
S22700  
amphiphysin - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C/Accession: S22700  
R;Lichte, B.; Veh, R.W.; Meyer, H.E.; Kilmann, M.W.  
EMBO J. 11, 2521-2530, 1992  
A;Title: Amphiphysin, a novel protein associated with synaptic vesicles.  
A;Reference number: S22700; MUID:92331604; PMID:1628617  
A;Accession: S22700  
A;Molecule type: mRNA  
A;Residues: 1-682 <LIC>  
A;Cross-references: EMBL:X60422; NID:g52842; PID:CAA42953.1; PID:g62843  
C;Superfamily: amphiphysin; RVS161 protein homology  
F;11-270/Domain: RVS161 protein homology <RVS>  
Query Match 93.7%; Score 89; DB 1; Length 682;  
Best Local Similarity 94.4%; Pred. No. 2.4e-06;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RIAXGRKLVYDSARHH 18  
DB 134 RIAXGRKLVYDSARHH 151  
RESULT 3  
S22400  
amphiphysin (clone 22-2) - human  
C:Species: Homo sapiens (man)  
C>Date: 12-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999  
C/Accession: S62400; I37166  
R;David, C.; Solimena, M.; de Camilli, P.  
PES Lett. 351, 73-79, 1994  
A;Title: Autoimmunity in Stiff-Man Syndrome with breast cancer is targeted to the C-term  
A;Reference number: S48686; MUID:94357284; PMID:8076697  
A;Accession: S62400  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-695 <DAV>  
A;Cross-references: EMBL:U07616; NID:g550449; PID:AAA21865.1; PID:g550450  
R;Yamamoto, R.; Li, X.; Winter, S.; Francke, U.; Kilmann, M.W.  
Hum. Mol. Genet. 4, 265-268, 1995  
A;Title: Primary structure of human amphiphysin, the dominant autoantigen of paraneoplas  
A;Reference number: I37166; MUID:95276740; PMID:7757077  
A;Accession: I37166  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-695 <RES>  
A;Cross-references: EMBL:X81438; NID:g662991; PID:CAA57197.1; PID:g662992  
C;Genetics:  
A;Gene: GDB:AMPH  
A;Map position: 7p14-7p13  
A;Cross-references: GDB:386990  
C;Superfamily: amphiphysin; RVS161 protein homology  
F;11-270/Domain: RVS161 protein homology <RVS>  
Query Match 93.7%; Score 89; DB 2; Length 695;  
Best Local Similarity 94.4%; Pred. No. 2.4e-06;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RIAXGRKLVYDSARHH 18  
DB 134 RIAXGRKLVYDSARHH 151

Db 134 RIAKGRKLVYDVSARHH 151

# RESULT 4

T22946

hypothetical protein F58G6.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000

C:Accession: T22946

R:Lloyd, C.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19641

A:Accession: T22946

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-461 <WIL>

A:Cross-references: EMBL:Z68217; PIDN:CAA92465.1; GSPDB:GN00022; CBSP:F58G6.1

A:Experimental source: clone F58G6

C:Genetics:

A:Gene: CBSP:F58G6.1

A:Map position: 4

A:Introns: 19/3; 57/3; 96/3; 130/3; 157/3; 186/3; 219/3; 271/3; 302/3; 335/3; 357/3

C:Superfamily: amphiphysin; RVSI61 protein homology

# Query Match

Best Local Similarity 72.6%; Score 69; DB 2; Length 461;

Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDVSARH 17

Db 130 KIEKGRKLVYDSAKN 146

# RESULT 5

F91120

hypothetical protein ECS3934 [imported] - Escherichia coli (strain O157:H7, substrain R1)

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: F91120

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F91120

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-130 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA837357.1; PID:gl3363407; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 050952

C:Genetics:

A:Gene: ECS3934

# Query Match

Best Local Similarity 52.1%; Score 49.5; DB 2; Length 130;

Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 5 RGRKLVYDVSARHH 18

Db 84 RGRGLIEGKGYDEARHH 100

# RESULT 6

F84259

hypothetical protein Vngl031c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: F84259

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.O.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: F84259

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-169 <STO>

A:Cross-references: GB:AE004437; NID:gl0580584; PIDN:AAG19442.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG1031C

C:Superfamily: conserved hypothetical protein MJ1050

# Query Match

Best Local Similarity 50.5%; Score 48; DB 2; Length 169;

Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 AKRGKLVYDVSARHH 18

Db 44 ADRGKLVYDVAVRDH 59

# RESULT 7

F84658

probable protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002

C:Accession: F84658

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: F84658

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-424 <STO>

A:Cross-references: GB:AE002093; NID:gl0375390; PIDN:AAC14522.1; GSPDB:GN00139

C:Genetics:

A:Gene: Atg26290

A:Map position: 2

C:Superfamily: kinase-related transforming protein; protein kinase homology

# Query Match

Best Local Similarity 49.5%; Score 47; DB 2; Length 424;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGRKLVYDVSARHH 18

Db 382 KEGKKFVDINKFRHH 396

# RESULT 8

TS1888

asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) [validated] - Arabidopsis thali

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 03-Nov-2000

C:Accession: TS1888

R:Lam, H.M.; Hsieh, M.H.; Coruzzi, G.

Plant J. 16, 345-353, 1998

A:Title: Reciprocal regulation of distinct asparagine synthetase genes by light and meta

A:Reference number: Z24846

A:Accession: TS1888

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-578 <LAM>

A:Cross-references: EMBL:AF095452; PIDN:AAC72836.1

C:Genetics:

A:Gene: ASN3

C:Function:

A:Description: EC 6.3.5.4 [validated, MUID:99097830]

C:Superfamily: asparagine synthase (glutamine-hydrolyzing)

C:Keywords: ligase

Query Match 49.5%; Score 47; DB 2; Length 578;  
 Best Local Similarity 53.3%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARHH 18  
 |||:|  
 DB 276 KAGREVADYLGTRHH 290

RESULT 9  
 T50812  
 asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) [similarity] - Arabidopsis thaliana  
 N;Alternate names: protein F18D22 10; protein T31P16.230  
 C;Species: Arabidopsis thaliana [mouse-ear cress]  
 C;Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 19-Jul-2002  
 C;Accession: T50812, T50028  
 R;Bavan, M.; Volckaert, G.; Grymonprez, B.; Voet, M.; Robben, J.; Bancroft, I.; Mewes, H.  
 submitted to the Protein Sequence Database, July 2000  
 A;Reference number: Z25241  
 A;Accession: T50812  
 A;Molecule type: DNA  
 A;Residues: 1-578 <BEV1>  
 A;Cross-references: EMBL:AL360334  
 A;Experimental source: cultivar Columbia; BAC clone F18D22  
 R;Bavan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; Smith, J.  
 submitted to the Protein Sequence Database, May 2000  
 A;Reference number: Z25027  
 A;Accession: T50028  
 A;Molecule type: DNA  
 A;Residues: 138-578 <BEV2>  
 A;Cross-references: EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.230  
 A;Experimental source: cultivar Columbia; BAC clone T31P16  
 C;Genetics:  
 A;Gene: ATSP:T31P16.230  
 A;Map position: 5  
 A;Introns: 27/2; 73/3; 105/3; 153/1; 185/1; 216/3; 270/3; 297/3; 371/3; 416/3; 443/3; 473/3; 500/3  
 C;Superfamily: asparagine synthase (glutamine-hydrolyzing)  
 C;Keywords: asparagine biosynthesis; ligase  
 F;2-578/Product: asparagine synthase (glutamine-hydrolyzing) #status predicted <MAT>  
 F;2/Active site: Cys #status predicted

Query Match 49.5%; Score 47; DB 2; Length 578;  
 Best Local Similarity 53.3%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARHH 18  
 |||:|  
 DB 276 KAGREVADYLGTRHH 290

RESULT 10  
 S40975  
 hypothetical protein F58A4.3 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Sep-1997  
 C;Accession: S40975  
 R;Berks, M.  
 submitted to the EMBL Data Library, February 1992  
 A;Reference number: S40975  
 A;Accession: S40975  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-288 <BER>  
 A;Cross-references: EMBL:Z22179; NID:G297956; PID:G297958  
 C;Genetics:  
 A;Introns: 31/2; 113/3; 214/3

Query Match 47.4%; Score 45; DB 2; Length 288;  
 Best Local Similarity 41.2%; Pred. No. 13;  
 Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 IAKRGKRLVDYDSARHH 18  
 :|:|:|  
 DB 14 LAQSDRKLADYLLAQPDTARRH 34

RESULT 11  
 T40661  
 yeast reduced viability upon starvation protein 161 homolog, implicated in cell growth a  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
 C;Accession: T40661  
 R;Lyne, M.; Rajadream, M.A.; Barrell, B.G.; Rieger, M.  
 submitted to the EMBL Data Library, November 1998  
 A;Reference number: Z21889

Query Match 45.8%; Score 43.5; DB 2; Length 282;  
 Best Local Similarity 47.6%; Pred. No. 23;  
 Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 2 IAKRGKRLVDYDSARHH 17  
 :|:|:|  
 DB 14 LAQSDRKLADYLLAQPDTARRH 34

RESULT 12  
 T40661  
 Probable transcription regulator yfhh [imported] - Salmonella enterica subsp. enterica s  
 C;Species: Salmonella enterica subsp. enterica serovar typhi  
 A;Note: this species has also been called Salmonella typhi  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C;Accession: AB0828  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.;  
 T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulie, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A;Reference number: AB0502; MUID:21534947; PMID:11677608  
 A;Accession: AB0828  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-282 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD02774.1; PID:gi6503784; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: yfhh  
 C;Superfamily: hypothetical protein ybbh

Query Match 46.3%; Score 44; DB 2; Length 225;  
 Best Local Similarity 63.2%; Pred. No. 15;  
 Matches 12; Conservative 3; Mismatches 2; Indels 2; Gaps 2;

QY 2 IAKRGKRLVDYDSARHH 18  
 |||:|  
 DB 65 IAEQGRLLVDFDSRLHH 83

RESULT 13  
 T40661  
 protein-glutamine gamma-glutamyltransferase  
 N;Alternate names: transglutaminase  
 C;Species: Chrysophrys major (red sea bream)  
 C;Date: 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 07-May-1999  
 C;Accession: PC2237  
 R;Yasueda, H.; Kumazawa, Y.; Motoki, M.  
 Biosci Biotechnol Biochem 58, 2041-2045, 1994  
 A;Title: Purification and characterization of a tissue-type transglutaminase from red sea  
 A;Reference number: PC2237; MUID:95078510; PMID:7765597  
 A;Accession: PC2237  
 A;Molecule type: protein  
 A;Residues: 1-225 <YAS>  
 A;Experimental source: liver  
 A;Comment: This enzyme catalyzes a Ca2+-dependent acyl-transfer reaction in which the ga  
 he acceptors.  
 C;Superfamily: protein-glutamine gamma-glutamyltransferase  
 C;Keywords: aminoacyltransferase

A:Accession: T40661  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-264 <LYN>  
 A:Cross-references: EMBL:AL034352; PIDN:CAA22181.1; GSPDB:GN00067; SPDB:SPBC725.09c  
 A:Experimental source: strain 972h-; cosmid c725  
 C:Genetics:  
 A:Gene: SPDB:SPBC725.09c  
 A:Map position: 2  
 A:Introns: 16/3  
 A:Superfamily: RVS161 protein; RVS161 protein homology  
 F:4-255/Domain: RVS161 protein homology <RVS>

Query Match 45.3%; Score 43; DB 2; Length 264;  
 Best Local Similarity 53.3%; Pred. No. 26;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IAKGRKLVYDSAR 16

Db 132 ITRNHHLLDHDAMR 146

# RESULT 14

T43000  
 RVS161 protein homolog - fission yeast (Schizosaccharomyces pombe) (fragment)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 02-Sep-2000  
 A:Accession: T43000  
 R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.  
 DNA Res. 4, 363-369, 1997  
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.  
 A:Reference number: 21723; MUID:98162722; PMID:9501391  
 A:Accession: T43000  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-270 <YOS>  
 A:Cross-references: EMBL:D89200; NID:G1749607; PIDN:BAAL3861.1; PID:G1749608  
 A:Experimental source: strain PR745  
 C:Superfamily: RVS161 protein; RVS161 protein homology  
 F:10-261/Domain: RVS161 protein homology <RVS>

Query Match 45.3%; Score 43; DB 2; Length 270;  
 Best Local Similarity 53.3%; Pred. No. 26;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IAKGRKLVYDSAR 16

Db 138 ITRNHHLLDHDAMR 152

# RESULT 15

S62579  
 Probable plasma membrane iron permease - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 11-Jan-2000  
 C:Accession: S62579; T38101  
 R:Murphy, L.; Niblett, D.; Harris, D.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: S62573  
 A:Accession: S62579  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-397 <MUR>  
 A:Cross-references: EMBL:Z67998; NID:G1067202; PIDN:CAA91954.1; PID:G1067209  
 R:Back, A.; Reinhardt, R.; Murphy, L.; Niblett, D.; Harris, D.; Barrell, B.G.; Rajandrea  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z21769  
 A:Accession: T38101  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-397 <MU2>  
 A:Cross-references: EMBL:Z67998; NID:G1067202; PIDN:CAA91954.1; GSPDB:GN00066; SPDB:SPAC  
 A:Experimental source: strain 972h-; cosmid c1f7

C:Genetics:  
 A:Gene: SPAC1F7.07c  
 A:Map position: 1R  
 C:Superfamily: conserved probable membrane protein VBR207w

Query Match 45.3%; Score 43; DB 2; Length 397;  
 Best Local Similarity 52.9%; Pred. No. 37;  
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IAKGRKLVYDSARHH 18

Db 323 IAKLGDKVVDLEAASSH 339

# RESULT 16

I55498  
 testicular dynamin - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 02-Feb-2001  
 C:Accession: I55498  
 R:Nakata, T.; Takemura, R.; Hirokawa, N.  
 J. Cell Sci. 105, 1-5, 1993

A:Title: A novel member of the dynamin family of GTP-binding proteins is expressed speci  
 A:Reference number: I55498; MUID:93366923; PMID:8360266  
 A:Accession: I55498  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-848 <RES>  
 A:Cross-references: GB:D14076; NID:G391871; PIDN:BAAO3161.1; PID:G391872  
 C:Superfamily: human dynamin II; pleckstrin repeat homology  
 C:Keywords: alternative splicing; GTP binding; membrane trafficking; nucleotide binding;  
 F:38-45/Region: nucleotide-binding motif A (P-loop)  
 F:132-137/Region: nucleotide-binding motif B  
 F:514-619/Domain: pleckstrin repeat homology <PLK>

Query Match 45.3%; Score 43; DB 2; Length 848;  
 Best Local Similarity 46.2%; Pred. No. 75;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 RGRKLVYDSARH 17

Db 87 KGKKFTDFDVRH 99

# RESULT 17

B69497  
 conserved hypothetical protein AF1979 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 12-Jun-2003  
 C:Accession: B69497  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: B69497  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-252 <KLE>  
 A:Cross-references: GB:AE000966; GB:AE000782; NID:G2689289; PIDN:AA889275.1; PID:G264856  
 C:Superfamily: nucleotidyltransferase

Query Match 44.2%; Score 42; DB 2; Length 252;  
 Best Local Similarity 63.6%; Pred. No. 35;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 GRKLVYDSAR 16

Db 30 GRKLVYDSAR 40

## RESULT 18

70933  
 hypothetical protein RV0573c - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 15-Jun-2001  
 C:Accession: D70933  
 R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
 A: Accession: D70933  
 A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-463 <COL>  
 A: Cross-references: GB:AL021942; GB:AL123456; NID: G3242298; PIDN: CAA17444.1; PID: G290963  
 A: Experimental source: strain H37RV  
 C: Geneticks:  
 A: Gene: RV0573c  
 C: Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF0821

Query Match 44.2%; Score 42; DB 2; Length 463;  
 Best Local Similarity 50.0%; Pred. No. 62;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 AKRGKLVYDVSARHH 18  
 |||||:|:|:|  
 Db 148 AARGPVDVFGARRAH 163

## RESULT 19

32607  
 trifunctional enzyme HDE, peroxisomal - Yeast (Candida tropicalis)  
 N: Alternate names: trifunctional beta-oxidation enzyme HDE  
 N: Contains: 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35); 3-hydroxyacyl-CoA epimerase  
 C: Species: Candida tropicalis  
 C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 31-Mar-2000  
 C: Accession: S32607; JH0474; JTO350; S18126  
 R: Sloots, J.A.; Aitchison, J.D.; Rachubinski, R.A.  
 Gene 105, 129-134, 1991  
 A: Title: Glucose-responsive and oleic acid-responsive elements in the gene encoding the  
 A: Reference number: S32607; MUID: 92039009; PMID: 1937002  
 A: Accession: S32607  
 A: Status: nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-906 <AIT>  
 A: Cross-references: EMBL: X57854; NID: G2669; PIDN: CAA40989.1; PID: G2670  
 A: Note: nucleotide sequence submitted to the EMBL Data Library, February 1991  
 R: Aitchison, J.D.; Sloots, J.A.; Nuttley, W.M.; Rachubinski, R.A.  
 Gene 105, 135-136, 1991

A: Title: Sequence of the gene encoding Candida tropicalis peroxisomal trifunctional enzyme  
 A: Reference number: JH0474; MUID: 92039010; PMID: 1937003  
 A: Accession: JH0474  
 A: Status: translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-906 <AIT>  
 A: Cross-references: EMBL: X57854; NID: G2669; PIDN: CAA40989.1; PID: G2670  
 R: Nuttley, W.M.; Aitchison, J.D.; Rachubinski, R.A.  
 Gene 69, 171-180, 1988  
 A: Title: cDNA cloning and primary structure determination of the peroxisomal trifunctional  
 A: Reference number: JTO350; MUID: 89172062; PMID: 3267241  
 A: Accession: JTO350  
 A: Molecule type: mRNA  
 A: Residues: 1-539, S', 541-906 <NUT>  
 A: Cross-references: GB: M22765; NID: G965397; PIDN: AAA62847.1; PID: G695398  
 A: Note: the authors translated the codon TCC for residue 540 as Phe  
 C: Geneticks:  
 A: Gene: HDE  
 C: Superfamily: peroxisomal trifunctional enzyme HDE; short-chain alcohol dehydrogenase  
 C: Keywords: carbon-oxygen lyase; fatty acid beta-oxidation; hydro-lyase; isomerase; muid

F: 9-194/Domain: short-chain alcohol dehydrogenase homology <SAD1>  
 F: 323-498/Domain: short-chain alcohol dehydrogenase homology <SAD2>

Query Match 44.2%; Score 42; DB 1; Length 906;  
 Best Local Similarity 46.7%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDVSAR 16  
 |||||:|:|:|  
 Db 267 VAKRFSILYDDSR 281

## RESULT 20

F83264  
 hypothetical protein PA3063 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C: Species: Pseudomonas aeruginosa  
 C: Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C: Accession: F83264  
 R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.; Nature 406, 959-964, 2000  
 A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
 A: Reference number: A82950; MUID: 20437337; PMID: 10984043  
 A: Accession: F83264  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-1193 <STO>  
 A: Cross-references: GB: AE004730; GB: AE004091; NID: G9949154; PIDN: AAG06451.1; GSPDB: GN001  
 A: Experimental source: strain PA01  
 C: Geneticks:  
 A: Gene: PA3063

Query Match 44.2%; Score 42; DB 2; Length 1193;  
 Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 KLVYDVSARHH 18  
 |||||:|:|:|  
 Db 81 KLGFEGQARHH 91

## RESULT 21

A59294  
 skeletal myosin - nematode (Onchocerca volvulus)  
 N: Alternate names: major body wall myosin  
 C: Species: Onchocerca volvulus  
 C: Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000  
 C: Accession: A59294  
 R: Werner, C.; Rajan, T.V.; Mol. Biochem. Parasitol. 50, 255-260, 1992  
 A: Title: Comparison of the body wall myosin heavy chain sequences from Onchocerca volvulus  
 A: Reference number: A59294; MUID: 92158005; PMID: 1741012  
 A: Accession: A59294  
 A: Status: preliminary; not compared with conceptual translation  
 A: Molecule type: mRNA  
 A: Residues: 1-1957 <WER>  
 A: Cross-references: GB: M74066; NID: G159892; PIDN: AAA29420.1; PID: G159893  
 A: Experimental source: sex female  
 C: Superfamily: myosin heavy chain; myosin motor domain homology  
 F: 87-773/Domain: myosin motor domain homology <MMO>

Query Match 44.2%; Score 42; DB 2; Length 1957;  
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 AKRGKLVYDVSARHH 18  
 |||||:|:|:|  
 Db 1702 AERARKQADYDANEAH 1717

## RESULT 22

A55800



cucumis (EC 3.4.21.25) precursor - muskmelon  
C;Species: Cucumis melo (muskmelon)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A55900  
R;Yamagata, H.; Masuzawa, T.; Nagaoaka, Y.; Ohnishi, T.; Iwasaki, T.  
J. Biol. Chem. 269, 32725-32731, 1994  
A;Title: Cucumis, a serine protease from melon fruits, shares structural homology with  
P;616-731/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
P;615-616/Cleavage site: Thr-Gly (autolytic) #status predicted  
A;Accession: A55800  
A;Molecule type: mRNA  
A;Residues: 'LIAKSTT', 1-731 <YAM>  
A;Cross-references: EMBL:D32206  
A;Note: it is uncertain whether Met-1 is the initiator or whether translation is initiated  
C;Superfamily: subtilisin-like proteinase ag12; subtilisin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-110/Domain: signal sequence and amino-terminal propeptide (fragment) #status predicted  
F;111-615/Product: cucumis #status predicted <WAT>  
F;131-539/Domain: subtilisin homology #status atypical <SBT>  
P;616-731/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
P;615-616/Cleavage site: Thr-Gly (autolytic) #status predicted  
Query Match 43.7%; Score 41.5; DB 1; Length 731;  
Best Local Similarity 61.1%; Pred. No. 1.1e+02;  
Matches 11; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 2 IAKRGKLVYDSAR-HH 18  
DB 35 IVTMGRKLEDPDGAHLHH 52  
RESULT 23  
MNVSUH  
nonstructural protein NS - snowshoe hare virus  
C;Species: snowshoe hare virus  
C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 30-Sep-1993  
C;Accession: A04105  
R;Bishop, D.H.L.; Gould, K.G.; Akashi, H.; Clerx-van Haaster, C.M.  
Nucleic Acids Res. 10, 3703-3713, 1982  
A;Title: The complete sequence and coding content of snowshoe hare bunyavirus small (S)  
A;Reference number: A93429; MUID:82274210; PMID:7050911  
A;Accession: A04105  
A;Molecule type: genomic RNA  
A;Residues: 1-92 <BIS>  
A;Note: this virus is a member of the California encephalitis serogroup of mosquito-transmitted bunyaviruses  
C;Comment: The genome consists of a large (L) RNA, a medium (M) RNA, and a small (S) RNA  
C;Superfamily: bunyavirus nonstructural protein  
Query Match 43.2%; Score 41; DB 1; Length 92;  
Best Local Similarity 37.5%; Pred. No. 20;  
Matches 6; Conservative 8; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RIAKRGKLVYDSAR 16  
DB 48 RVSQGRQILNLESGR 63  
RESULT 24  
T13595  
hypothetical protein 80H7.10 - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C;Accession: T13595  
R;Benos, P.  
Submitted to the EMBL Data Library, April 1999  
A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
A;Reference number: Z17667  
A;Accession: T13595  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-196 <BEN>  
A;Cross-references: EMBL:AL031027; NID:e1313443; PID:e1427308; PIDN:CAB41536.1  
C;Genetics:  
A;Cross-references: FlyBase:FBgn0000481

A;Introns: 52/1; 96/3; 140/3  
A;Note: EG:80H7.10

Query Match 43.2%; Score 41; DB 2; Length 196;  
Best Local Similarity 64.3%; Pred. No. 41;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 AKRGKLVYDSAR 16  
DB 45 AMEGRKLEIDWAR 58

## RESULT 25

E70945  
hypothetical protein Rv2054 - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C;Accession: E70945  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A;Authors: Searles, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: E70945  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-237 <COL>  
A;Cross-references: GB:AL021899; GB:AL123456; NID:93242282; PIDN:CAA17268.1; PID:93242282  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: Rv2054  
C;Superfamily: carboxymethylglutaminase  
Query Match 43.2%; Score 41; DB 2; Length 237;  
Best Local Similarity 52.9%; Pred. No. 49;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDSARHH 18  
DB 79 LTKRGALDDILAARDH 95

## RESULT 26

T25732  
hypothetical protein F25B4.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C;Accession: T25732  
R;Gattung, S.  
Submitted to the EMBL Data Library, July 1996  
A;Description: The sequence of C. elegans cosmid F25B4.  
A;Reference number: Z20076  
A;Accession: T25732  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-402 <GAT>  
A;Cross-references: EMBL:U64842; PIDN:AAB37080.1; GSPDB:GN00023; CESP:F25B4.1  
A;Experimental source: strain Bristol N2; clone F25B4  
C;Genetics:  
A;Gene: CESP:F25B4.1  
A;Map position: 5  
A;Introns: 25/2; 87/2; 289/1  
C;Superfamily: aminomethyltransferase  
Query Match 43.2%; Score 41; DB 2; Length 402;  
Best Local Similarity 43.8%; Pred. No. 79;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDSARH 17  
DB 288 VAKRRRTLDFFGAH 303

RESULT 27  
C71055  
Probable RNA methyltransferase - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: C71055  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekir  
M.; Ohtoku, Y.; Funahashi, T.; Tanaka, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: C71055  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-407 <KAW>  
A:Cross-references: GB:AP000005; NID:G3236132; PIDN:BAA30237.1; PID:G3257554  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH137  
C:Superfamily: hypothetical protein HI0333

Query Match 43.2%; Score 41; DB 2; Length 407;  
Best Local Similarity 61.5%; Pred. No. 80;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IAKGRKLVVDYDS 14  
:|||||:|:  
Db 289 LAKRGFKVGVGFD 301

RESULT 28  
D89057  
protein K09H11.1 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: D89057  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: D89057  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-974 <STO>  
A:Cross-references: GB:chr\_V; PIDN:AAB52261.1; PID:G1938424; GSPDB:GN00023; CESP:K09H11.  
C:Genetics:  
A:Gene: K09H11.1  
A:Map position: 5

Query Match 43.2%; Score 41; DB 2; Length 974;  
Best Local Similarity 72.7%; Pred. No. 1.8e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 GRKLVVDYDSAR 16  
:|||||:|:  
Db 862 GRKLVQVDSLR 872

RESULT 29  
T50914  
hypothetical membrane protein [imported] - Rubrivivax gelatinosus  
C:Species: Rubrivivax gelatinosus  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50914  
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.  
submitted to the EMBL Data Library, November 1999  
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn  
A:Reference number: Z25270

A:Accession: T50914  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1001 <NAG>  
A:Cross-references: EMBL:AB034704; PIDN:BAA94067.1  
A:Experimental source: strain Ili44  
C:Genetics:  
A:Note: ORF1001

Query Match 43.2%; Score 41; DB 2; Length 1001;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RIAXGRKLVVDYDSAR 16  
:|||||:|:  
Db 772 RFARQGRLFSRDDAR 787

RESULT 30  
H65033  
hypothetical protein in purL-dpj intergenic region - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Mar-2002  
C:Accession: H65033  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: H65033  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-306 <BLAT>  
A:Cross-references: GB:AE000342; GB:U00096; NID:G1788907; PIDN:AAC75614.1; PID:G1788913.  
A:Experimental source: strain K-12, substrain MGL655  
C:Genetics:  
A:Gene: yfjH  
C:Superfamily: hypothetical protein yfjH

Query Match 42.6%; Score 40.5; DB 1; Length 306;  
Best Local Similarity 42.9%; Pred. No. 74;  
Matches 9; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

Qy 2 IAKRGKLVVDY----DSARH 17  
:|||||:|:  
Db 38 LAQSDKKLADYLLQLQDPARH 58

RESULT 31  
C91057  
hypothetical protein ECs3427 [imported] - Escherichia coli (strain O157:H7, substrain R)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 17-May-2002  
C:Accession: C91057  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: C91057  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-306 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA036850.1; PID:G13362898; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECs3427  
C:Superfamily: hypothetical protein yfjH

Query Match 42.6%; Score 40.5; DB 2; Length 306;  
Best Local Similarity 42.9%; Pred. No. 74;  
Matches 9; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

Qy 2 IAKGRKLVY-----DSARH 17  
:|: :||| | | | :|||  
Db 38 LAOSDKKLADYLLLOPDTARH 58

## RESULT 32

G85901  
hypothetical protein yfhH [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C/Species: Escherichia coli  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002  
C/Accession: G85901  
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A/Reference number: A85480; MUID:21074935; PMID:11206551  
A/Accession: G85901  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-306 <STC>  
A/Cross-references: GB:AE005174; NID:g12516974; PIDN:AA657675.1; GSPDB:GN00145; UWGSP:Z38  
A/Experimental source: strain O157:H7, substrain EDL933  
C/Genetics:  
C/Gene: yfhH  
C/Superfamily: hypothetical protein ybbH

Query Match 42.6%; Score 40.5; DB 2; Length 306;  
Best Local Similarity 42.9%; Pred. No. 74;  
Matches 9; Conservative 4; Mismatches 3; Indels

Qy 2 IAKRGRKLVDP-----DSARH 17  
          ::: :||| | | | | |  
Db 38 LAOSDKKLADYLLLOPDTARH 58

RESULT 33

A71288  
probable methylenetetrahydrofolate dehydrogenase (fold) - syphilis spirochete  
C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C/Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 03-Dec-1999  
C/Accession: A71288  
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin-  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo-  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A/Reference number: A71250; MUID:98332770; PMID:9665876  
A/Accession: A71288  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-315 <COL>  
A/Cross-references: GB:AE001245; GB:AE000520; NID:g33233033; PIDN:AAC65701.1; PID:g3323303  
A/Experimental source: strain Nichols  
C/Genetics:  
A/Gene: TP0732  
C/Superfamily: methylenetetrahydrofolate dehydrogenase (NAD+); methylenetetrahydrofolate  
F5-288/Domain: methylenetetrahydrofolate dehydrogenase (NAD+) homology <MTPD>

Query Match 42.6%; Score 40.5; DB 2; Length 315;  
Best Local Similarity 50.0%; Pred. No. 76;  
Matches 9; Conservative 3; Mismatches 3; Indels

Qy	3	AKRGEKL---	VDYDSARH	17
Db	242	APRGRLCGDV	DFDAVAH	259

RESULT 34

AC2896  
 oxidoreductase Atu2604 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C;Accession: AC2896

R.; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kucyavin, T.; Levy, R.; Li, M.; McClellan, K.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J. E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MIID:21608550; PMID:11743193

A;Accession number: AF012604  
A;Accession: AC2896  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-393 <KUR>  
A;Cross-references: GB:AF012604  
A;Experimental source: *Salmonella typhimurium*  
C;Genetics:  
A;Gene: *Atu2604*

A:Map position:

Query Match 42.6%; Score 40.5; DB 2; Length 393;  
Best Local Similarity 45.5%; Pred. No. 93;  
Matches 10: Conservative 2; Mismatches 3: Indels

QY 3 AKGRKLV D-----YDSARH 17  
: : : : :  
DQ 361 ARHGRKVDLKPWVEEYDLALH 38

RESULT 35

Probable oxidoreductase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
 F97871  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C:Accession: F97671  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gol  
 A.; Liu, F.; Wollan, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacteriu

A:Reference number:

A:Accession: J16711  
A>Status: preliminary  
A/Molecule type: DNA  
A/Molecule type: DNA  
A/Residues: 1-408 <KUR>  
A/Cross-references: GB:AE007869; FIDN:AAK8327.1; PID:gi15157803; GSPDB:GN00169  
C:Genetics:

A;Gene: AGR\_C\_4719  
A;Map position: circular chromosome

Query Match	42.6%	Score 40.5;	DB 2;	Length 408;
Best Local Similarity	45.5%	Pred. No. 96;		
Matches 10; Conservative	2;	Mismatches 3;	Indels 7;	Gaps 1;

QY 3 AKRGRKLV-----YDSARH 17  
| : ||| : || |  
Db 376 ARHGRKVVDLKPVTVEEYDLALH 397

## RESIT.T 36

S77016 sensory transduction histidine kinase sl10790 - *Synechocystis* sp. (strain PCC 6803)  
N;Alternate names: protein sl10790  
C;Species: *Synechocystis* sp.  
A;Variety: PCC 6803

C:\msd2\...  
#sequence revision 25-Apr-1997 #text change 02-Sep-2000

C/Accession: S77016  
R/Kaneoko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z

A;Reference number: S74322; MUID:97061201; PMID:9905231  
A;Accession: S77016  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA



A;Residues: 1-651 <BEV>  
 A;Cross-references: EMBL:AL353995; GSPDB:GN00063; ATSP:F12B17.120  
 A;Experimental source: cultivar Columbia; BAC clone F12B17  
 C;Genetics:  
 A;Gene: ATSP:F12B17.120  
 A;Map position: 5  
 C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology

Query Match 42.1%; Score 40; DB 2; Length 651;  
 Best Local Similarity 61.5%; Pred. No. 1.8e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GRKLVYDYSARH 18  
 Db 193 GRVLIFYDSARN 195

RESULT 42  
 S56146  
 GCN20 protein - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: protein YFR009w  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 14-Oct-1995 #sequence\_revision 16-Feb-1996 #text\_change 02-Feb-2001  
 C;Accession: S56146; S56264  
 R;Vazquez de Aldana, C.R.; Marton, M.J.; Hinnebusch, A.G.  
 EMBO J. 14, 3184-3199, 1995  
 A;Title: GCN20, a novel ATP binding cassette protein, and GCN1 reside in a complex that  
 A;Reference number: S56146; MUID:95347344; PMID:7621831  
 A;Accession: S56146  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: DNA  
 A;Residues: 1-752 <VAZ>  
 A;Cross-references: EMBL:U19971; NID:G643478; PIDN:AAA75444.1; PID:G643479  
 R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasano  
 submitted to the EMBL Data Library, May 1995  
 A;Description: Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cerevisiae*  
 A;Reference number: S56264  
 A;Accession: S56264  
 A;Molecule type: DNA  
 A;Residues: 1-752 <MUR>  
 A;Cross-references: EMBL:D50617; NID:G836685; PIDN:BAA09248.1; PID:dl009889; PID:G836764  
 C;Genetics:  
 A;Gene: SGD:GCN20  
 A;Cross-references: SGD:S0001905; MIPS:YFR009w  
 A;Map position: 6R  
 C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

P;215-440/Domain: ATP-binding cassette homology <ABC1>  
 F;232-239/Region: nucleotide-binding motif A (P-loop)  
 F;548-723/Domain: ATP-binding cassette homology <ABC2>  
 F;565-572/Region: nucleotide-binding motif A (P-loop)

Query Match 42.1%; Score 40; DB 2; Length 752;  
 Best Local Similarity 37.5%; Pred. No. 2e+02;  
 Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RIARGRKLVYDYSAR 16  
 Db 152 KVAKRNKKFKYKYEASK 167

RESULT 43  
 T18860  
 Hypothetical protein C02C6.1 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
 C;Accession: T18860  
 R;Swinnburne, J.  
 Submitted to the EMBL Data Library, August 1996  
 A;Reference number: Z19032  
 A;Accession: T18860  
 A;Status: preliminary; translated from GB/EMBL/DBDJ  
 A;Molecule type: DNA

A;Residues: 1-830 <WIL>  
 A;Cross-references: EMBL:Z79596; PIDN:CAB01857.1; GSPDB:GN00028; CESP:C02C6.1  
 A;Experimental source: clone C02C6  
 C;Genetics:  
 A;Gene: CESP:C02C6.1  
 A;Map position: X  
 C;Superfamily: human dynamin II; pleckstrin repeat homology

Query Match 42.1%; Score 40; DB 2; Length 830;  
 Best Local Similarity 46.2%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KRGRKLVYDYSAR 16  
 Db 88 KKGHRFVDFDAVR 100

RESULT 44  
 H64367  
 dolichyl-phosphate mannose synthase - *Methanococcus jannaschii*  
 C;Species: *Methanococcus jannaschii*  
 C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C;Accession: H64367  
 R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.;  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
 A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
 A;Reference number: A64300; MUID:96337999; PMID:8688087  
 A;Accession: H64367  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-229 <BUL>  
 A;Cross-references: GB:U67504; GB:L77117; NID:gl591248; PIDN:AAB98538.1; PID:gl592295; T  
 C;Genetics:  
 A;Map position: REV482606-481917  
 C;Superfamily: *Methanobacterium thermoautotrophicum* dolichyl-phosphate mannose synthase

Query Match 41.6%; Score 39.5; DB 2; Length 229;  
 Best Local Similarity 42.1%; Pred. No. 82;  
 Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

Qy 1 RIARGRKLV-DYDSARH 18  
 Db 185 OLAKKGAKTVEYVSVEYH 203

RESULT 45  
 T31023  
 conserved hypothetical protein 108 - *Sulfolobus* sp. plasmid pNOB8  
 C;Species: *Sulfolobus* sp.  
 C;Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
 C;Accession: T31023  
 R;She, Q.; Phan, H.; Garrett, R.A.; Alberte, S.V.; Stedman, K.M.; Zillig, W.  
 Extremophiles 2, 417-425, 1998  
 A;Title: Genetic profile of pNOB8 from *Sulfolobus*: the first conjugative plasmid from ar  
 A;Reference number: Z20959; MUID:99044580; PMID:9827331  
 A;Accession: T31023  
 A;Status: preliminary; translated from GB/EMBL/DBDJ  
 A;Molecule type: DNA  
 A;Residues: 1-97 <SHE>  
 A;Cross-references: EMBL:AJ010405; NID:e1351926; PID:e1351945; PIDN:CAA09129.1  
 A;Experimental source: strain NOB8H2  
 C;Genetics:  
 A;Genome: plasmid pNOB8

Query Match 41.1%; Score 39; DB 2; Length 97;  
 Best Local Similarity 50.0%; Pred. No. 45;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RIARGRKLVYD 12

T01186 protein kinase (EC 2.7.1.37) tousled 4 - maize (fragment)  
N;Alternate names: tousled-like kinase 4  
C;Species: Zea mays (maize)

Qy 4 KRGRKLVYDSARHH 18  
D<sup>b</sup> 125 KHGYDSNDYSSVRRHH 139

T01186 protein kinase (EC 2.7.1.37) tousled 4 - maize (fragment)  
N;Alternate names: tousled-like kinase 4  
C;Species: Zea mays (maize)

A:Gene: ATSP:TI6L24.50

A:Map position: 3  
A:introns: 47/3; 80/3; 131/2

Query Match 41.1%; Score 39; DB 2; Length 259;  
Best Local Similarity 85.7%; Pred. NO. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 YDSARHH 18

|||:||||  
Db 227 YDSRRHH 233

Search completed: March 4, 2004, 17:47:00  
Job time : 20.6774 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:32:31 ; Search time 12.1935 Seconds

(without alignments)  
76.865 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_138\_155

Perfect score: 95

Sequence: 1 RIAGRGRKLVYDSARHH 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	95	100.0	588	1 BIN1 MOUSE
2	95	100.0	588	1 BIN1 RAT
3	95	100.0	593	1 BIN1 HUMAN
4	89	93.7	682	1 AMPH CHICK
5	89	93.7	683	1 AMPH RAT
6	89	93.7	695	1 AMPH HUMAN
7	45	47.4	288	1 YMH3 CAEL
8	43	45.3	204	1 COAE RALSO
9	43	45.3	264	1 HOBI SCHPO
10	43	45.3	397	1 YAK7 SCHPO
11	43	45.3	775	1 YAK7 MOUSE
12	43	45.3	848	1 DYN3 RAT
13	42	44.2	130	1 YC35 SKECO
14	42	44.2	906	1 FOX2 CANTR
15	41	43.2	92	1 VNST BUNSH
16	41	43.2	321	1 MIAA RHIL0
17	41	43.2	454	1 WCM SHIFL
18	40.5	42.6	282	1 YFHE ECOLI
19	40	42.1	406	1 SNX6 HUMAN
20	40	42.1	461	1 KLB2 ECOLI
21	40	42.1	752	1 GC20 YEAST
22	40	42.1	830	1 DYN1 CAEL
23	39.5	41.6	229	1 Y544 METJA
24	39	41.1	356	1 LEUD TALSO
25	39	41.1	512	1 MSL3 DROME
26	39	41.1	776	1 ANR5 HUMAN
27	39	41.1	918	1 QAL5 NEUCR
28	39	41.1	980	1 POLG LIV
29	39	41.1	981	1 YM48 YEAST
30	39	41.1	2894	1 YD96 METJA
31	38	40.0	203	1 YPT1 VOLCA
32	38	40.0	359	1 FLTI EUGGL
33	38	40.0	413	1 ARG7 BRAJA

34	38	40.0	431	1 PDA6 RAT	Q63081 rattus norv
35	38	40.0	440	1 PDA6 HUMAN	Q5084 homo sapien
36	38	40.0	464	1 WCM ECOLI	P1244 escherichia
37	38	40.0	478	1 YH33 CHROV	Q7x94 chromobacte
38	38	40.0	504	1 YAI8 TREPA	Q83981 treponema p
39	38	40.0	642	1 PHR NEUCR	P27526 neurospora
40	38	40.0	651	1 NAD0 THEBR	P32382 thermocanaer
41	38	40.0	715	1 STC2 STAAU	P17855 staphylococ
42	38	40.0	957	1 SECA MYCSM	P15333 mycobacteri
43	38	40.0	1434	1 NOS1 HUMAN	P29475 homo sapien
44	38	40.0	1435	1 NOS1 RABIT	O19132 oryctolagus
45	38	40.0	1518	1 KKK1 YEAST	P34244 saccharomyc
46	38	40.0	2418	1 SPCA HUMAN	P02549 homo sapien
47	38	40.0	4473	1 PLE1 CRIGR	Q9155 cricetus
48	38	40.0	4684	1 PLE1 HUMAN	Q15149 homo sapien
49	37.5	39.5	347	1 VCOM ADEL2	P36717 human adeno
50	37	38.9	88	1 RS15 HAEIN	P43389 haemophilus
51	37	38.9	91	1 Y022 ARCFU	Q02113 archaeoglob
52	37	38.9	128	1 Y453 ARCFU	Q29796 archaeoglob
53	37	38.9	145	1 RL13 STACA	Q00990 staphylococ
54	37	38.9	147	1 YV59 CAEL	P50439 caenorhabdi
55	37	38.9	150	1 CKS1 YEAST	P20486 saccharomyc
56	37	38.9	163	1 GVH2 HALNI	Q9ht6 halobacteri
57	37	38.9	163	1 GVH2 HALNI	P33961 halobacteri
58	37	38.9	278	1 PRXH EPMD2	O64252 mycobacteri
59	37	38.9	315	1 LXD1 PHOLE	P21309 photobacter
60	37	38.9	331	1 MAN1 MOUSE	Q9wu40 mus musculu
61	37	38.9	338	1 AROB PYRAB	Q9v1h9 pyrococcus
62	37	38.9	365	1 ISEG PROST	P72241 providencia
63	37	38.9	376	1 MCL1 CITR	Q4845 citrobacter
64	37	38.9	439	1 PDA6 MESAU	P38660 mesocricetu
65	37	38.9	525	1 PDA2 HUMAN	Q13087 homo sapien
66	37	38.9	532	1 YKF3 CAEL	P34266 caenorhabdi
67	37	38.9	535	1 GCP METJA	Q58530 methanococc
68	37	38.9	541	1 Y537 SYN3	P54148 synechocyst
69	37	38.9	585	1 ASNS MAIZE	P49094 zea mays (m
70	37	38.9	624	1 PPS1 CAVEO	O54820 c.bifunctio
71	37	38.9	662	1 DCHS HUMAN	P19113 homo sapien
72	37	38.9	697	1 SYGB BUCBP	P59573 buchnera ap
73	37	38.9	761	1 FIXI RHILV	Q33533 rhizobium l
74	37	38.9	808	1 SYFB BACHD	Q9K896 bacillus ha
75	37	38.9	892	1 SYA PYRAE	Q8zsv6 pyrobaculum
76	37	38.9	900	1 SYA AERPE	Q8y9x3 aeropyrum p
77	37	38.9	925	1 SYA METPA	Q8tu79 methanosarc
78	37	38.9	926	1 SYA METMA	Q8pwu0 methanosarc
79	37	38.9	1139	1 INAI CAEL	Q03600 caenorhabdi
80	37	38.9	1671	1 DPOL PYRKO	P77933 pyrococcus
81	36.5	38.4	70	1 RPOZ THETN	Q8r9s7 thermocanaer
82	36.5	38.4	107	1 L14K TORMA	Q81499 torpedo mar
83	36.5	38.4	269	1 ML22 LACUC	P50178 lactococcus
84	36.5	38.4	358	1 VCOM ADE40	P48753 human adeno
85	36.5	38.4	431	1 SYS NEIMA	Q9jt77 neisseria m
86	36.5	38.4	431	1 SYS NEIMA	Q9jt77 neisseria m
87	36.5	38.4	638	1 PDA4 MOUSE	P08003 mus musculu
88	36.5	38.4	643	1 PDA4 RAT	P38659 rattus norv
89	36	37.9	51	1 RL39 METKA	Q8tuv3 methanopyru
90	36	37.9	88	1 RS15 ECOLI	P02371 escherichia
91	36	37.9	88	1 RS15 ECOLI	Q8xif9 salmonella
92	36	37.9	114	1 YZ11 METJA	Q60273 methanococc
93	36	37.9	164	1 NUSB CHLFE	Q8kbs4 chlorobium
94	36	37.9	182	1 KYV SULAC	P35029 sulfolobus
95	36	37.9	188	1 KCV THEAC	Q9ht3 thermoplasm
96	36	37.9	211	1 HIS5 RHOSH	O33565 rhodobacter
97	36	37.9	231	1 RL1 BACHD	Q8KGS bacillus ha
98	36	37.9	249	1 YF57 METJA	Q58952 methanococc
99	36	37.9	260	1 EMD RAT	Q63190 rattus norv
100	36	37.9	299	1 DH31 CAEL	Q9xt10 caenorhabdi
101	36	37.9	299	1 NODP RHITR	P52995 rhizobium t
102	36	37.9	324	1 PARB TREPA	Q83295 treponema p
103	36	37.9	354	1 DCUP XYLTT	Q87dvo xyella fas
104	36	37.9	368	1 HISX PYRAE	Q8zy17 pyrobaculum
105	36	37.9	398	1 CGB2 BOVIN	O77689 bos taurus
106	36	37.9	398	1 CGB2 MOUSE	P30276 mus musculu





Query Match 100.0%; Score 95; DB 1; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 139 RIAGRGKLVYDSARHH 156

RESULT 2  
 BIN1\_RAT  
 ID BIN1\_RAT  
 AC 00839; STANDARD; PRT; 588 AA.

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Myc box dependent interacting protein 1 (Bridging integrator 1)  
 DE (Amphiphysin-like protein) (Amphiphysin II)  
 GN BIN1 OR AMPHL OR AMPH2.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

SEQUENCE FROM N.A., SUBUNIT, AND ALTERNATIVE SPLICING.  
 RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex, and kidney;  
 RX MEDLINE=98009145; PubMed=9348539;  
 RA Wigge P., Kohler K., Hunt S., Doyle C., McMahon H.T.;  
 RT "Amphiphysin heterodimers: potential role in clathrin-mediated  
 endocytosis";  
 RL Mol. Biol. Cell 8:2003-2015(1997).  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM AMPH2-1).  
 RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex;  
 RX MEDLINE=97424383; PubMed=9280305;  
 RA McMahon H.T., Wigge P., Smith C.;  
 RT "Clathrin interacts specifically with amphiphysin and is displaced by  
 dynamin.";  
 RL FEBS Lett. 413:319-322(1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 506-588, FUNCTION, AND  
 RP INTERACTION WITH DNMI.  
 RX MEDLINE=98409533; PubMed=9736607;  
 RA Owen D.J., Wigge P., Vallis Y., Moore J.D.A., Evans P.R.,  
 RA McMahon H.T.;  
 RT "Crystal structure of the amphiphysin-2 SH3 domain and its role in the  
 RP prevention of dynamin ring formation.";  
 RL EMBO J. 17:5273-5285(1998).

-!- FUNCTION: May be involved in regulation of synaptic vesicle  
 endocytosis. May act as a tumor suppressor and inhibits  
 malignant cell transformation.  
 CC -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).  
 CC Binds to SYN1 and DNMI through its SH3 domain, and to clathrin  
 CC through a region outside of the SH3 domain. Also binds to alpha-  
 CC adaptin. Interacts with the N-terminal transactivation domain of  
 CC MYC in a manner requiring the integrity of the conserved MYC box  
 CC regions 1 and 2.  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=6;  
 CC Name=AMPH2-1;  
 CC IsoId=008839-1; Sequence=Displayed;  
 CC Name=AMPH2-2;  
 CC IsoId=008839-2; Sequence=VSP\_000260;  
 CC Name=AMPH2-3;  
 CC IsoId=008839-3; Sequence=VSP\_000258;  
 CC Name=AMPH2-4;  
 CC IsoId=008839-4; Sequence=VSP\_000256, VSP\_000257;  
 CC Name=AMPH2-5;  
 CC IsoId=008839-5; Sequence=VSP\_000259;  
 CC Name=AMPH2-6;  
 CC IsoId=008839-6; Sequence=VSP\_000256, VSP\_000259;  
 CC -!- TISSUE SPECIFICITY: Isoform AMPH2-1 is expressed in brain.

concentrated at nerve terminals. Isoform AMPH2-2 is widely  
 expressed.  
 CC -!- PTM: Phosphorylated by protein kinase C.  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Y13380; CAA73807.1; -  
 PDB; 1BB9; 17-JUN-98.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR..  
 DR InterPro; IPR004148; BAR\_dom.  
 DR InterPro; IPR001452; SH3..  
 DR Pfam; PF03114; SH3; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR PRODOM; PD000066; SH3; 1.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW Alternative splicing; SH3 domain; Coiled coil; Endocytosis;  
 KW Anti-oncogene; Differentiation; Phosphorylation; 3D-structure.  
 FT DOMAIN 15 42  
 FT DOMAIN 193 274  
 FT DOMAIN 379 422  
 FT DOMAIN 515 588  
 FT VARSPLIC 173 205  
 FT VARSPLIC 253 588  
 FT VARSPLIC 335 588  
 FT VARSPLIC 335 482  
 FT VARSPLIC 423 460  
 FT VARSPLIC 588 AA; 64533 MW; 164AC90E09547F1A CRC64;  
 SQ SEQUENCE 588 AA; 64533 MW; 164AC90E09547F1A CRC64;  
 Query Match 100.0%; Score 95; DB 1; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
 139 RIAGRGKLVYDSARHH 156

RESULT 3  
 BIN1\_HUMAN  
 ID BIN1\_HUMAN STANDARD; PRT; 593 AA.  
 AC 000499; 000297; 000545; 043867; 060552; 060553; 060554; 060555;  
 AC 075514; 075515; 075516; 075517; 075518; 092944; 099688;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Myc box dependent interacting protein 1 (Bridging integrator 1)  
 DE (Amphiphysin-like protein) (Amphiphysin II) (Box-dependent myc-  
 DE interacting protein-1).  
 GN BIN1 OR AMPHL  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1] SEQUENCE FROM N.A. (ISOFORM IIA).  
 RP TISSUE=Brain;  
 RC MEDLINE=97341217; PubMed=9195986;  
 RX Ramjaun A.R., Micheva K.D., Bouchalet I., McPherson P.S.;  
 RA "Identification and characterization of a nerve terminal-enriched  
 RT amphiphysin isoform."  
 RL J. Biol. Chem. 272:16700-16706(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS IIA AND BIN1).  
 RP TISSUE=Brain, and Skeletal muscle;  
 RC MEDLINE=97327761; PubMed=9182667;  
 RX Butler M.H., David C., Ochoa G.-C., Freyberg Z., Daniell L., Grabs D.,  
 RA Crenona O., De Camilli P.;  
 RT "Amphiphysin II (SH3P9; BIN1), a member of the amphiphysin/Rvs family,  
 is concentrated in the cortical cytomatrix of axon initial segments  
 and nodes of ranvier in brain and around T tubules in skeletal  
 muscle."  
 RL J. Cell Biol. 137:1355-1367(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM BIN1).  
 RP TISSUE=Skeletal muscle;  
 RC MEDLINE=96376973; PubMed=8782822;  
 RX Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
 RA "BIN1 is a novel Myc-interacting protein with features of a tumour  
 RT suppressor."  
 RL Nat. Genet. 14:69-76(1996).  
 RN [4]  
 RP REVISIONS TO N-TERMINUS.  
 RA Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS IIB; IIC1; IIC2 AND IID).  
 RP TISSUE=Brain;  
 RC MEDLINE=96264340; PubMed=9603201;  
 RX Ramjaun A.R., McPherson P.S.;  
 RA "Multiple amphiphysin II splice variants display differential clathrin  
 RT binding: identification of two distinct clathrin-binding sites."  
 RL J. Neurochem. 70:2369-2376(1998).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS IIC2 AND IIC3).  
 RP TISSUE=Brain;  
 RC MEDLINE=97366618; PubMed=9223448;  
 RX Tsutsui K., Maeda Y., Tsutsui K., Seki S., Tokunaga A.;  
 RA "cDNA cloning of a novel amphiphysin isoform and tissue-specific  
 RT expression of its multiple splice variants."  
 RL Biochem. Biophys. Res. Commun. 236:178-183(1997).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORMS IIC3; IIC3; BIN1-10-13; BIN1-13 AND  
 RP BIN1+12A).  
 RP TISSUE=Fibroblast;  
 RC MEDLINE=98058932; PubMed=9395479;  
 RX Wechsler-Reya R.J., Sakamuro D., Zhang J., DuHadaway J.,  
 RA Prendergast G.C.;  
 RT "Structural analysis of the human BIN1 gene. Evidence for tissue-  
 RT specific transcriptional regulation and alternate RNA splicing."  
 RL J. Biol. Chem. 272:31453-31458(1997).  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORM IIC2).  
 RA Zhang J., Du W., Wechsler-Reya R.J., DuHadaway J., Sakamuro D.,  
 RA Prendergast G.C.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE OF 133-593 FROM N.A.  
 RP TISSUE=Brain;  
 RC Yu W., Gibbs R.A.;  
 RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [10]  
 RP CHARACTERIZATION.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=98078712; PubMed=9418903;  
 RA Wechsler-Reya R.J., Elliott K.J., Prendergast G.C.;  
 RT "A role for the putative tumor suppressor Bin1 in muscle cell

RT differentiation."  
 RL Mol. Cell. Biol. 18:566-575(1998).  
 CC -!- FUNCTION: May be involved in regulation of synaptic vesicle  
 CC endocytosis. May act as a tumor suppressor and inhibits  
 CC malignant cell transformation.  
 CC -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).  
 CC Binds to SYNJ1 and DNMI through its SH3 domain, and to clathrin  
 CC through a region outside of the SH3 domain. Also binds to alpha-  
 CC adaptin. Interacts with the N-terminal transactivation domain of  
 CC MYC in a manner requiring the integrity of the conserved MYC box  
 CC regions 1 and 2.  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Isoform IIA is  
 CC found in the cytoplasm while isoform BIN1 is nuclear.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=11;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=IIA; Sequence=Displayed;  
 CC IsoId=O00499-1; Sequence=Displayed;  
 CC Name=IIB; Sequence=VSP\_000246; VSP\_000252;  
 CC IsoId=O00499-2; Sequence=VSP\_000246; VSP\_000252;  
 CC Name=IIC1; Sequence=VSP\_000249;  
 CC IsoId=O00499-3; Sequence=VSP\_000249;  
 CC Name=IIC2; Sequence=VSP\_000246; VSP\_000249;  
 CC IsoId=O00499-4; Sequence=VSP\_000246; VSP\_000249;  
 CC Name=IID; Sequence=VSP\_000248;  
 CC IsoId=O00499-5; Sequence=VSP\_000248;  
 CC Name=II2; Sequence=VSP\_000246; VSP\_000253;  
 CC IsoId=O00499-6; Sequence=VSP\_000246; VSP\_000253;  
 CC Name=II3; Sequence=VSP\_000246; VSP\_000250;  
 CC IsoId=O00499-7; Sequence=VSP\_000246; VSP\_000250;  
 CC Name=BIN1; Sequence=VSP\_000246; VSP\_000247; VSP\_000250;  
 CC IsoId=O00499-8; Sequence=VSP\_000246; VSP\_000247; VSP\_000250;  
 CC Name=BIN1-10-13; Sequence=VSP\_000246; VSP\_000251;  
 CC IsoId=O00499-9; Sequence=VSP\_000246; VSP\_000251;  
 CC Name=BIN1-13; Sequence=VSP\_000246; VSP\_000251;  
 CC IsoId=O00499-10; Sequence=VSP\_000246; VSP\_000247; VSP\_000251;  
 CC Name=BIN1+12A; Sequence=VSP\_000246; VSP\_000247; VSP\_000253;  
 CC IsoId=O00499-11; Sequence=VSP\_000246; VSP\_000247; VSP\_000253;  
 CC -!- TISSUE SPECIFICITY: ISOFORM IIA IS EXPRESSED ONLY IN THE BRAIN  
 CC WHERE IT IS CONCENTRATED IN AXON INITIAL SEGMENTS AND NODES OF  
 CC RANVIER. ISOFORM BIN1 IS WIDELY EXPRESSED WITH HIGHEST EXPRESSION  
 CC IN SKELETAL MUSCLE.  
 CC -!- PTM: Phosphorylated by protein kinase C (By similarity).  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF004015; AAC51345.1; -  
 CC EMBL; AF070576; AAC28646.1; -  
 CC EMBL; U68485; AAC17461.1; -  
 CC EMBL; AF001383; AAC61363.1; -  
 CC EMBL; AF043898; AAC39710.1; -  
 CC EMBL; AF043899; AAC39711.1; -  
 CC EMBL; AF043900; AAC39712.1; -  
 CC EMBL; AF043901; AAC39713.1; -  
 CC EMBL; U87558; AAB63263.1; -  
 CC EMBL; AF068914; AAC24126.1; -  
 CC EMBL; AF068915; AAC24127.1; -  
 CC EMBL; AF068916; AAC24128.1; -  
 CC EMBL; AF068917; AAC23750.1; -  
 CC EMBL; AF068918; AAC23751.1; -  
 CC EMBL; U84004; AAC23440.1; -  
 CC EMBL; U83999; AAC23440.1; JOINED.  
 CC EMBL; U84001; AAC23440.1; JOINED.  
 CC EMBL; U84002; AAC23440.1; JOINED.  
 CC EMBL; U84003; AAC23440.1; JOINED.

```

DR EMBL; U84004; AAC23441.1; ALT INIT.
DR EMBL; U83999; AAC23441.1; JOINED.
DR EMBL; U84001; AAC23441.1; JOINED.
DR EMBL; U84002; AAC23441.1; JOINED.
DR EMBL; U84003; AAC23441.1; JOINED.
DR PIR; JCS593; JCS593.
DR TRANSFAC; T03490; -.
DR Genew; HGNC:1052; BIN1.
DR MIM; 601248; -.
DR GO; GO:0015629; C:actin cytoskeleton; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR GO; GO:0006899; P:nonsselective vesicle transport; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR004148; BAR dom.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD003208; Amphiphysin_1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
DR CytoSkelton; SH3 domain; Coiled coil.
KW DOMAIN 10 84 COILED COIL (POTENTIAL).
FT DOMAIN 144 191 COILED COIL (POTENTIAL).
FT DOMAIN 609 682 SH3.
SQ SEQUENCE 682 AA; 75205 MW; 61617F494F38EB20 CRC64;

Query Match 93.7%; Score 89; DB 1; Length 682;
Best Local Similarity 94.4%; Pred. No. 2.3e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RIAKGRKLVYDSARHH 18
Db 134 RIAKGRKLVYDSARHH 151

RESULT 5
AMPH_RAT
ID AMPH_RAT STANDARD; PRT; 683 AA.
AC Q08838;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amphiphysin.
GN AMPH OR AMPH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain cortex;
RX MEDLINE=98009145; PubMed=9348539;
RA Wigge P., Kohler K., Vallis Y., Owen D., Hunt S.P., McMahon H.T.;
RT "Amphiphysin heterodimers: potential role in clathrin-mediated
endocytosis."
RL Mol. Biol. Cell 8:2003-2015 (1997).
CC -1- FUNCTION: May participate in mechanisms of regulated exocytosis in
synapses and certain endocrine cell types. May control the
properties of the membrane associated cytoskeleton (By

```

CC similarity).

CC -!- SUBUNIT: Heterodimer with BIN1. Binds SH3GLB1 (By similarity).

CC -!- SUBCELLULAR LOCATION: Peripheral membrane protein associated with

CC the cytoplasmic surface of synaptic vesicles (By similarity).

CC -!- SIMILARITY: Contains 1 BAR domain.

CC -!- SIMILARITY: Contains 1 SH3 domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; Y13381; CAA73808.1; -

CC InterPro; IPR003005; Amphiphysin.

CC InterPro; IPR003017; Amphiphysin\_1.

CC InterPro; IPR006632; BAR.

CC InterPro; IPR001452; SH3.

CC Pfam; PF03114; BAR; 1.

CC Pfam; PF00018; SH3; 1.

CC PRINTS; PR01251; AMPHIPHYSIN.

CC PRINTS; PR00452; SH3DOMAIN.

CC ProDom; PD003208; Amphiphysin\_1; 1.

CC ProDom; PD000066; SH3; 1.

CC SMART; SM00721; BAR; 1.

CC SMART; SM00326; SH3; 1.

CC PROSITE; PS00002; SH3; 1.

CC CytoSkelton; SH3 domain; Coiled coil.

CC FT DOMAIN 10 83 COILED COIL (POTENTIAL).

CC FT DOMAIN 144 191 COILED COIL (POTENTIAL).

CC FT DOMAIN 610 683 SH3.

CC SEQUENCE 683 AA; 74877 MW; 7FEA4A9E5A1F6631 CRC64;

Query Match 93.7%; Score 89; DB 1; Length 683;

Best Local Similarity 94.4%; Pred. No. 2.3e-07;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAGKGRKLVYDSARHH 18

DB 134 RIAGKSRKLVYDSARHH 151

RESULT 6

AMPH HUMAN STANDARD; PRT; 695 AA.

AC P49418; O43538;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Amphiphysin.

GN AMPH OR AMPH1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]\_TaxID=9606;

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Cerebellum;

RX MEDLINE=94357284; PubMed=8076697;

RA David C., Solimena M., de Camilli P.;

RT "Autoimmunity in stiff-man syndrome with breast cancer is targeted to

RT the C-terminal region of human amphiphysin, a protein similar to the

RT yeast proteins, Rvs167 and Rvs161.";

RL FEBS Lett. 351:73-79(1994).

RN [2]\_

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Brain;

RX MEDLINE=95276740; PubMed=7757077;

RA Yamamoto R., Li X., Winter S., Francke U., Kilimann M.W.;

RT "Primary structure of human amphiphysin, the dominant autoantigen of

RT paraneoplastic stiff-man syndrome, and mapping of its gene (AMPH) to

RT Chromosome 7p13-p14.";

RL Hum. Mol. Genet. 4:265-268(1995).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Brain;

RX MEDLINE=98174372; PubMed=9513187;

RA Floyd S.R., Butler M.H., Cremona O., David C., Freyberg Z., Zhang X.,

RA Solimena M., Tokunaga A., Ishizu H., Tsutsui K., De Camilli P.V.;

RT "Expression of amphiphysin 1, an autoantigen of paraneoplastic

RT neurological syndromes, in breast cancer.";

RL Mol. Med. 4:29-39(1998).

CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in

CC synapses and certain endocrine cell types. May control the

CC properties of the membrane associated cytoskeleton.

CC -!- SUBUNIT: Heterodimer of AMPH and AMPH1 (By similarity).

CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC SURFACE OF

CC SYNAPTIC VESICLES.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1; Synonyms=128 kDa;

CC IsoId=P49418-1; Sequence=Displayed;

CC Name=2; Synonyms=108 kDa;

CC IsoId=P49418-2; Sequence=VSP\_000245;

CC -!- TISSUE SPECIFICITY: Neurons, certain endocrine cell types and

CC spermatocytes.

CC -!- DISEASE: Patients with stiff-man syndrome, a rare disease of the

CC central nervous system characterized by progressive rigidity of

CC the body musculature with superimposed painful spasms, have

CC antibodies against AMPH.

CC -!- SIMILARITY: Contains 1 BAR domain.

CC -!- SIMILARITY: Contains 1 SH3 domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U07616; AAA21865.1; -

CC EMBL; X81438; CAA57197.1; -

CC EMBL; AF034996; AAC02977.1; -

CC PIR; S62400; S62400.

CC PDB; 1KY7; 12-JUN-02.

CC Genew; HGNC:471; AMPH.

CC MIM; 600418; -

CC GO; GO:0015629; C:actin cytoskeleton; TAS.

CC GO; GO:0008021; C:synaptic vesicle; TAS.

CC GO; GO:0006897; P:endocytosis; TAS.

CC GO; GO:0007268; P:synaptic transmission; TAS.

CC InterPro; IPR003005; Amphiphysin.

CC InterPro; IPR003017; Amphiphysin\_1.

CC InterPro; IPR006632; BAR.

CC InterPro; IPR001452; SH3.

CC Pfam; PF03114; BAR; 1.

CC Pfam; PF00018; SH3; 1.

CC PRINTS; PR01251; AMPHIPHYSIN.

CC PRINTS; PR00452; SH3DOMAIN.

CC ProDom; PD003208; Amphiphysin\_1; 1.

CC ProDom; PD000066; SH3; 1.

CC SMART; SM00721; BAR; 1.

CC SMART; SM00326; SH3; 1.

CC PROSITE; PS00002; SH3; 1.

CC CytoSkelton; SH3 domain; Coiled coil; Alternative splicing;

CC 3D-structure.

CC FT DOMAIN 10 83 COILED COIL (POTENTIAL).

CC FT DOMAIN 144 191 COILED COIL (POTENTIAL).

CC FT DOMAIN 622 695 SH3.

CC VARSPLIC 425 466 Missing (in isoform 2).

CC /FTId=VSP\_000245.

QY 2 IAKRGRKLVDYDSARHH 18

OS Schizosaccharomyces pombe (Fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RX MEDLINE=21293007; PubMed=11274158;  
 RA Routhier E.L., Burn T.C., Abbaszade I., Summers M., Albright C.F.,  
 RA Prendergast G.C.; complements the F-actin localization defects caused by  
 RT loss of Hob3p, the fission yeast homolog of Kvs1c1p.";  
 RL J. Biol. Chem. 276:21670-21677(2001).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21849401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehmann W., Wambutt R., Purnelle B.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomycetes pombe."  
 RL Nature 415:871-880(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PR745;  
 RX MEDLINE=98162722; PubMed=9501991;  
 RA Yoshikawa S., Kato K., Nakai K., Okayama H., Nojima H.;  
 RA "Identification of open reading frames in Schizosaccharomycetes pombe  
 RT CDNA's."  
 RL DNA Res. 4:363-369(1997).  
 CC -!- FUNCTION: Involved in cytokinesis and septation where it has a  
 CC role in the localization of F-actin.  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AF275638; AAF86459.1; --  
 CC EMBL; AL034352; CAA22181.1; --  
 CC EMBL; DB9200; BAAL3861.1; ALT\_INIT.  
 CC PIR; T40661; T40661.  
 CC GeneDB Spombe; SPBC725.09c; --  
 CC InterPro; IPR003005; Amphiphysin.  
 CC InterPro; IPR006632; BAR.  
 CC Pfam; PF03114; BAR; 1.  
 CC PRINTS; PR01251; AMPHIPHYSIN.

DR SMART; SM00721; BAR; 1.  
 KW Septation; Cytoskeleton; Coiled coil.  
 FT DOMAIN 25 65 COILED COIL (POTENTIAL).  
 FT DOMAIN 165 187 COILED COIL (POTENTIAL).  
 FT CONFLICT 93 93 Q -> L (IN REF. 3).  
 FT CONFLICT 96 96 E -> K (IN REF. 3).  
 SQ SEQUENCE 264 AA; 30094 MW; 845388582305AD4D CRC64;  
 Query Match 45.3%; Score 43; DB 1; Length 264;  
 Best Local Similarity 53.3%; Pred. No. 6.2;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 Qy 2 IAKRGKRLVDYDSAR 16  
 Db 132 ITRNKKLHDHMR 146  
 RESULT 10  
 YAK7 SCHPO STANDARD; PRT; 397 AA.  
 ID YAK7 SCHPO  
 AC Q09919; 1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein ClF7.07c in chromosome I.  
 GN SPAC187.07C.  
 OS Schizosaccharomycetes pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21849401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehmann W., Wambutt R., Purnelle B.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomycetes pombe."  
 RL Nature 415:871-880(2002).  
 RN [1]  
 RP SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: Belongs to the FTR1 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; Z67998; CAA91954.1; --

```

DR PIR: S62579; S62579.
DR GeneDB Spombe; SPAC1F7.07c; -.
DR InterPro: IPR004923; FTR1.
DR Pfam: PF03239; FTR1; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 92 112 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 292 312 POTENTIAL.
SQ SEQUENCE 397 AA; 44314 MW; 1D98E41AD1FD708F CRC64;

Query Match 45.3%; Score 43; DB 1; Length 397;
Best Local Similarity 52.9%; Pred. No. 9.6;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IAKRGKLVVDYSARHH 18
||| |::| |::|
Db 323 IAKLGKVDVLEMASSH 339

RESULT 11
ANRS_MOUSE
ID ANRS_MOUSE STANDARD; PRT; 775 AA.
AC Q9D2J7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ankyrin repeat domain protein 5.
GN ANKRD5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojoberi T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pasole G., Quackenbush J.,
RA Schirf L.M., Staehli P., Suzuki R., Tomika M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli K., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
CC -!- SIMILARITY: Contains 8 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AK019547; BAB31791.1; -.
CC HSSP: P80144; 2MYO.
CC MGD: MGI:2441685; ANK.
CC InterPro: IPR002110; ANK.
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00023; ank; 10.

```

```

DR PRINTS; PRO1415; ANKYRIN.
DR SMART; SMO0248; ANK; 10.
DR PROSITE; PS00088; ANK_REPEAT; 6.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Repeat. 76
FT REPEAT 47 76 ANK 1.
FT REPEAT 184 213 ANK 2.
FT REPEAT 217 246 ANK 3.
FT REPEAT 250 279 ANK 4.
FT REPEAT 524 553 ANK 5.
FT REPEAT 557 586 ANK 6.
FT REPEAT 590 619 ANK 7.
FT REPEAT 623 652 ANK 8.
SQ SEQUENCE 775 AA; 86903 MW; 51C11E1BC5151F00 CRC64;

Query Match 45.3%; Score 43; DB 1; Length 775;
Best Local Similarity 47.1%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 IAKRGKLVVDYSARHH 18
||| |::| |::|
Db 204 ILRGGEVNDNRHH 220

RESULT 12
DYN3_RAT
ID DYN3_RAT STANDARD; PRT; 848 AA.
AC Q08877;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dynamin 3 (EC 3.6.1.50) (Dynamin, testicular) (T-dynamin).
GN DNM3 OR DYN3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=93366923; PubMed=8360266;
RA Nakata T., Takamura R., Hirokawa N.;
RA "A novel member of the dynamin family of GTP-binding proteins is
RA expressed specifically in the testis.";
RT J. Cell Sci. 105:1-5(1993).
RL CC -!- FUNCTION: Microtubule-associated force-producing protein involved
CC in producing microtubule bundles and able to bind and hydrolyze
CC GTP. Most probably involved in vesicular trafficking processes, in
CC particular endocytosis.
CC -!- CATALYTIC ACTIVITY: GTP + H(2)O = GDP + phosphate.
CC -!- SUBCELLULAR LOCATION: Microtubule-associated.
CC -!- TISSUE SPECIFICITY: Expressed in germ-cell-depleted testis,
CC indicating its expression in Sertoli cells.
CC -!- SIMILARITY: Belongs to the dynamin family.
CC -!- SIMILARITY: Contains 1 PH domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D14076; BAA03161.1; -.
CC PIR: I55498; I55498.
CC HSSP: Q05193; 2DYN.
CC InterPro: IPR001401; Dynamin.
CC InterPro: IPR000375; Dynamin_central.
CC InterPro: IPR003130; GED.
CC InterPro: IPR001849; PH.
CC Pfam: PF00350; dynamin_1.
CC Pfam: PF01031; dynamin_2; 1.

```



DR Pfam; PF00212; GED; 1.  
 DR PRINTS; PF00169; PH; 1.  
 DR PROSITE; PS00195; DYNAMIN.  
 DR SMART; SM00053; DYNC; 1.  
 DR SMART; SM00302; GED; 1.  
 DR SMART; SM00233; PH; 1.  
 DR PROSITE; PS00410; DYNAMIN; 1.  
 DR PROSITE; PS00003; PH\_DOMAIN; 1.  
 KW Hydrolyase; Motor protein; GTP-binding; Microtubule; Multigene family;  
 KW Endocytosis.  
 FT NP\_BIND 38 45 GTP (BY SIMILARITY).  
 FT NP\_BIND 136 140 GTP (BY SIMILARITY).  
 FT NP\_BIND 205 208 GTP (BY SIMILARITY).  
 FT DOMAIN 515 621 PH.  
 FT SEQUENCE 848 AA; 95595 MW; 8023655FCFC685F6 CRC64;  
 Query Match 45.3%; Score 43; DB 1; Length 848;  
 Best Local Similarity 46.2%; Pred. No. 22;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 5 RGRKLVYDSARH 17  
 : : : : :  
 DB 87 KGKFTDFEVRH 99  
 : : : : :  
 RESULT 13  
 YC35\_SKECO STANDARD; PRT; 130 AA.  
 AC 096805;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical 15.1 kDa protein ycf35.  
 GN YCF35.  
 OS Skeletonema costatum (Marine centric diatom).  
 OG Chloroplast.  
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;  
 OC Thalassiosirophyceae; Thalassiosirales; Skeletonemataceae;  
 OC Skeletonema.  
 ON NCBI\_TaxID=2843;  
 OX 1;  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=NIES-323;  
 RX MEDLINE=20189624; PubMed=10727090;  
 RA Tada N., Shibata S., Otsuka S., Namba K., Oyazui H.;  
 RT "Comparison of gene arrangements of chloroplasts between two centric  
 RT diatoms, Skeletonema costatum and Odontella sinensis.";  
 RL DNA Seq. 10:343-347(1999).  
 CC -1- SIMILARITY: Belongs to the ycf35 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AJ132264; CA10626.1; -  
 DR Chloroplast; Hypothetical protein.  
 SQ SEQUENCE 130 AA; 15110 MW; 9A86185636915727 CRC64;  
 Query Match 44.2%; Score 42; DB 1; Length 130;  
 Best Local Similarity 40.0%; Pred. No. 4.3;  
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 KRGKLVYDSARH 18  
 : : : : :  
 DB 29 KREKKAIDANNKXH 43  
 : : : : :  
 RESULT 14  
 FOX2\_CANTR

ID AC  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Peroxisomal hydratase-dehydrogenase-epimerase (HDE) (Multifunctional  
 DE beta-oxidation protein) (MPP) [Includes: 2-enoyl-CoA hydratase  
 DE (EC 4.2.1.-); D-3-hydroxyacyl CoA dehydrogenase (EC 1.1.1.-)].  
 OS Candida tropicalis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5492;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 20336 / PK233;  
 RX MEDLINE=89172062; PubMed=3267241;  
 RA Nuttley W.M., Aitchison J.D., Rachubinski R.A.;  
 RT "cDNA cloning and primary structure determination of the peroxisomal  
 RT trifunctional enzyme hydratase-dehydrogenase-epimerase from the yeast  
 RT Candida tropicalis PK233.";  
 RL Gene 69:171-180(1988).  
 RN 2;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 20336 / PK233;  
 RX MEDLINE=92039009; PubMed=1937002;  
 RA Aitchison J.D., Sloots J.A., Nuttley W.M., Rachubinski R.A.;  
 RT "Glucose-responsive and oleic acid-responsive elements in the gene  
 RT encoding the peroxisomal trifunctional enzyme of Candida  
 RT tropicalis.";  
 RL Gene 105:129-134(1991).  
 RN 3;  
 RP SIMILARITY TO SHORT CHAIN DEHYDROGENASES OF N-TERMINAL DOMAIN.  
 RX MEDLINE=90367890; PubMed=2394320;  
 RA Baker M.E.;  
 RT "A common ancestor for Candida tropicalis and dehydrogenases that  
 RT synthesize antibiotics and steroids.";  
 RL FASEB J. 4:3028-3032(1990).  
 CC -1- FUNCTION: Second trifunctional enzyme acting on the beta-oxidation  
 CC pathway for fatty acids, possessing hydratase-dehydrogenase-  
 CC epimerase activities. Converts trans-2-enoyl-CoA via D-3-  
 CC hydroxyacyl-CoA to 3-ketoacyl-CoA.  
 CC -1- PATHWAY: Fatty acid beta-oxidation cycle; step 2.  
 CC -1- PATHWAY: Fatty acid beta-oxidation cycle; step 3.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -1- INDUCTION: By growth on N-alkanes or fatty acids.  
 CC -1- DOMAIN: Contains two SDR domains.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SDR) family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M22765; AAA62847.1; -  
 DR EMBL; X57854; CAA40989.1; -  
 DR PIR; S32607; S32607.  
 DR HSP; O70351; 1E3S.  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR002539; MacC dehydratas.  
 DR Pfam; PF00106; adh short; 2.  
 DR Pfam; PF01575; MacC dehydratas; 1.  
 DR PRINTS; PR00080; SDRFAMILY.  
 DR PROSITE; PS00061; ADH\_SHORT; 2.  
 DR Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD;  
 KW Lyase; Isomerase; Peroxisome; Repeat.  
 FT DOMAIN 5 228 SHORT-CHAIN DEHYDROGENASE LIKE 1.  
 FT DOMAIN 319 532 SHORT-CHAIN DEHYDROGENASE LIKE 2.  
 FT SITE 904 906 MICROBODY TARGETING SIGNAL (POTENTIAL).

FT NP BIND 12 36 NAD (BY SIMILARITY).  
 FT ACT SITE 163 163 BY SIMILARITY.  
 FT NP BIND 326 350 NAD (BY SIMILARITY).  
 FT ACT SITE 467 467 BY SIMILARITY.  
 FT CONFLICT 540 540 F -> S (IN REF. 1).  
 SQ SEQUENCE 906 AA; 99469 MW; 65AB88C8671C967B CRC64;  
 Query Match 44.2%; Score 42; DB 1; Length 906;  
 Best Local Similarity 46.7%; Pred. No. 35;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 IAKRGKLVYDYSAR 16  
 DB 267 VAKRFSEILDYDSR 281  
 RESULT 15  
 VNST\_BUNSH  
 ID\_VNST\_BUNSH STANDARD; PRT; 92 AA.  
 AC P03514;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Nonstructural protein NS-S.  
 DE Bunyavirus snowshoe hare.  
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.  
 OX NCBI\_TaxID=11580;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=82274210; PubMed=7050911;  
 RX Clerkx-Van Haaster C.M., Akashi H., Auperin D.D., Bishop D.H.L.;  
 RA "The complete sequence and coding content of snowshoe hare bunyavirus small (S) viral RNA species."  
 RT Nucleic Acids Res. 10:3703-3713(1982).  
 RL [2]  
 RN SEQUENCE OF 1-47 FROM N.A.  
 RP MEDLINE=82216937; PubMed=7086954;  
 RX Clerkx-Van Haaster C.M., Akashi H., Auperin D.D., Bishop D.H.L.;  
 RA "Nucleotide sequence analyses and predicted coding of bunyavirus genome RNA species."  
 RT J. Virol. 41:119-128(1982).  
 RL [1]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J02390; -; NOT ANNOTATED\_CDS.  
 CC EMBL; J02393; -; NOT ANNOTATED\_CDS.  
 CC PIR; A04105; MNVUSH.  
 CC InterPro; IPR000797; Bunya\_Nss.  
 CC Pfam; PF01104; Bunya\_NS-S\_1.  
 CC ProDom; PD002170; Bunya\_Nss; 1.  
 KW Nonstructural protein.  
 SQ SEQUENCE 92 AA; 10490 MW; B5CCF0C9FD4FA8C1 CRC64;  
 Query Match 43.2%; Score 41; DB 1; Length 92;  
 Best Local Similarity 37.5%; Pred. No. 4.3;  
 Matches 6; Conservative 8; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RIAKRGKLVYDYSAR 16  
 DB 48 RVSQRGRILNLSGR 63  
 RESULT 16  
 MIAA\_RHILO  
 ID\_MIAA\_RHILO STANDARD; PRT; 321 AA.  
 AC Q98KJ4;  
 DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase) (IPTase) (IPPT).  
 DE MIAA OR MLL1448.  
 GN Rhizobium loti (Mesorhizobium loti).  
 OS Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099; PubMed=11214968;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;  
 RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti."  
 RT DNA Res. 7:331-338(2000).  
 RL Mesorhizobium loti.  
 CC -!- FUNCTION: Catalyzes the first step in the biosynthesis of 2-methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS2I[6A]) adjacent to the anticodon of several tRNA species (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.  
 CC -!- SIMILARITY: Belongs to the Ipp transferase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF002997; BAB48820.1; -;  
 CC HAMAP; MF\_00185; -; 1.  
 CC InterPro; IPR002627; IPPT.  
 CC Pfam; PF01715; IPPT; 1.  
 CC ProDom; PD004674; IPPT; 1.  
 CC TIGRFAMs; TIGR00174; miaA; 1.  
 KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;  
 FT NP BIND 24 31 ATP (POTENTIAL).  
 SQ SEQUENCE 321 AA; 35410 MW; D7FA3E4E7F4EB12F CRC64;  
 Query Match 43.2%; Score 41; DB 1; Length 321;  
 Best Local Similarity 57.1%; Pred. No. 17;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 AKRGKLVYDYSAR 16  
 DB 193 AERGRPLIDRQTAR 206  
 RESULT 17  
 WCAM\_SHIFL  
 ID\_WCAM\_SHIFL STANDARD; PRT; 464 AA.  
 AC P37775;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Colanic acid biosynthesis protein wcam.  
 KW WCAM OR SF2106 OR S2229.  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Shigella.  
 OX NCBI\_TaxID=623;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=301 / Serotype 2a;

```

RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Xu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
[3]
RN SEQUENCE OF 156-464 FROM N.A.
RP STRAIN=PS577 / Serotype 2a;
RC MEDLINE=3411953; PubMed=7507920;
RA Morona R., Mavris M., Fallarino A., Manning P.A.;
RT "Characterization of the rfc region of Shigella flexneri."
RL J. Bacteriol. 176:733-747(1994).
CC -1- PATHWAY: Slime polysaccharide colanic acid biosynthesis.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A3015226; A043645.1; -
DR EMBL; A3016985; A041747.1; -
DR EMBL; X71970; CA50765.1; -
KW Lipopolysaccharide biosynthesis; Complete proteome.
FT CONFLICT 395 395 E -> D (IN REF. 3).
FT CONFLICT 398 398 V -> L (IN REF. 3).
FT CONFLICT 455 464 INFRLEPERE -> Y (IN REF. 3).
SQ SEQUENCE 464 AA; 51191 MW; 0EC3BA7001C7D43D CRC64;
Query Match 43.2%; Score 41; DB 1; Length 464;
Best Local Similarity 41.2%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 2 IAKRGKLVYDSARH 18
Db 432 VNERGSSVDIDRINH 448
RESULT 18
YFHH_ECOLI STANDARD; PRT; 282 AA.
AC P37767; P76588;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative HTH-type transcriptional regulator yfhh.
GN YFHH OR B2561.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RA Nashimoto H., Saito N.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=K12 / M51655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1233-1238(1997).
[3]
RN SEQUENCE OF 1-91 FROM N.A.
RP STRAIN=NWL37;
RC MEDLINE=92292954; PubMed=1602968;
RA Poulsen L.K., Larsen N.W., Molin S., Andersson P.;
RT "Analysis of an Escherichia coli mutant strain resistant to the cell-
RT killing function encoded by the gef gene family."
RL Mol. Microbiol. 6:895-905(1992).
[4]
RN IDENTIFICATION.
RP MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
RT bacterial genome."
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -1- SIMILARITY: Contains 1 SIS domain.
CC -1- SIMILARITY: Contains 1 HTH RPIR-type DNA-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D54044; BAA10911.1; ALT_INIT.
DR EMBL; U36841; AAY79823.1; ALT_INIT.
DR EMBL; A5000342; AAC75614.1; ALT_INIT.
DR EMBL; X72336; -; NOT ANNOTATED_CDS.
DR EMBL; EG12308; YFHH.
KW EcoGene; EGI2308; YFHH.
DR InterPro; IPR000281; HTH_RPIR.
DR InterPro; IPR001347; SIS.
DR Pfam; PF01418; HTH_6; 1.
DR Pfam; PF01380; SIS_1; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
FT DOMAIN 1 110 HTH_RPIR-TYPE.
FT DOMAIN 126 282 SIS.
FT DNA_BIND 37 56 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 282 AA; 30707 MW; FFOAE2A0BF4E3C48 CRC64;
Query Match 42.6%; Score 40.5; DB 1; Length 282;
Best Local Similarity 42.9%; Pred. No. 18;
Matches 9; Conservative 4; Mismatches 3; Indels 5; Gaps 1;
QY 2 IAKRGKLVYDSARH 17
Db 14 LAQSDKKLADYLLQPDYARH 34
RESULT 19
SNX6_HUMAN STANDARD; PRT; 406 AA.
AC Q9UNF7; Q9Y449;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sorting nexin 6 (TRAF4-associated factor 2).
GN SNX6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21378165; PubMed=11485546;
RA Teasdale R.D., Loci D., Houghton F., Karlsson L., Gleeson P.A.;
RT "A large family of endosome-localized proteins related to sorting
RT nexin 1.";
RL Biochem. J. 358:7-16(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Toji S., Yano M., Kobayashi A., Tamai K.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=21276432; PubMed=11279102;
RA Parks W.T., Frank D.B., Huff C., Haft C.R., Martin J., Meng X.,
RA de Caestecker M.P., McNally J.G., Reddi A., Taylor S.I., Roberts A.B.,
RA Wang T., Lechleider R.J.;
RT "Sorting nexin 6, a novel SNX, interacts with the transforming growth
RT factor-beta family of receptor serine-threonine kinases.";
RL J. Biol. Chem. 276:19332-19339(2001).
RN [4]
RP FUNCTION: May be involved in several stages of intracellular
RP trafficking.
CC -|- SUBUNIT: Interacts with TGF-beta receptors.
CC -|- SIMILARITY: Belongs to the sorting nexin family.
CC -|- SIMILARITY: Contains 1 phox homology (PX) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF121856; RAD27829.1; -
CC EMBL; U83194; AAD24202.1; ALT_INIT.
CC Genew; HGNC:14970; SNX6.
CC MIM; 606098; -
CC GO; GO:0005622; C:intracellular; NAS.
CC GO; GO:0005478; F:intracellular transporter activity; NAS.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0006886; P:intracellular protein transport; NAS.
CC GO; GO:0007175; P:negative regulation of EGF receptor activity; NAS.
CC InterPro; IPR001683; PX.
CC Pfam; PF00787; PX; 1.
CC PROSITE; PS50195; PX; 1.
CC TRANSPORT; Protein transport.
CC DOMAIN 26 173
CC SEQUENCE 406 AA; 46648 MW; E3659DB19C59E1BB CRC64;
Query Match 42.1%; Score 40; DB 1; Length 406;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 KRGRKLVVDNSA 15
DB 312 RRSRLVDYENA 323
RESULT 20
KLB2_ECOLI
ID KLB2_ECOLI STANDARD; PRT; 461 AA.
AC PS2605;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE KLB2 protein.
GN KLB2.
OS Escherichia coli.
OC Plasmid Incp-alpha RK2.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.

```

```

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94282811; PubMed=8014987;
RA Pansegrau W., Lanka E., Barth P.T., Figurski D.H., Quiney D.G.,
RA Haas D., Hellinski D.R., Schwab H., Stanisich V.A., Thomas C.M.;
RT "Complete nucleotide sequence of Birmingham Incp alpha plasmids.
RT Compilation and comparative analysis.";
RL J. Mol. Biol. 239:623-663(1994).
RN [2]
RP SEQUENCE OF 1-237 FROM N.A.
RA Larsen M.H., Figurski D.H.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP DOMAIN: CONTAINS A DNA-BINDING REGION JOINED BY A SHORT VARIABLE
RP SEGMENT TO A REGION SIMILAR TO E. COLI KORA AND TRBA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L27758; -; NOT ANNOTATED_CDS.
CC EMBL; U05773; AAA57449.1; -
CC Plasmid; DNA-binding; Transcription regulation.
CC SEQUENCE 461 AA; 51165 MW; 10EBFB2F2E00BD81 CRC64;
Query Match 42.1%; Score 40; DB 1; Length 461;
Best Local Similarity 53.8%; Pred. No. 37;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 RIARKEGRKLVVDYD 13
DB 432 RINRNGRKVCYD 444
RESULT 21
GC20_YEAST
ID GC20_YEAST STANDARD; PRT; 752 AA.
AC P43535;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE GCN20 protein.
GN GCN20 OR YFR009W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95347344; PubMed=7670463;
RA Vazquez de Aldana C.R., Marton M.J., Hinnebusch A.G.;
RT "GCN20, a novel ATP binding cassette protein, and GCN1 reside in a
RT complex that mediates activation of the eIF-2 alpha kinase GCN2 in
RT amino acid-starved cells.";
RL EMBO J. 14:3184-3199(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=5288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP FUNCTION: Positive effector of the eIF-2-alpha kinase activity of
RP GCN2.
CC -|- SUBUNIT: Component of a heteromeric complex that includes GCN1 and
CC GCN20.

```

```

CC -1- SIMILARITY: Belongs to the ABC transporter family, EF3 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19971; AAA75444.1; -
DR EMBL; D50617; BAA09248.1; -
DR PIR; S56146; S56146.
DR GERM; S0001905; GCM20.
DR GO; GO:0005830; Cytosolic ribosome (sensu Eukarya); IDA.
DR GO; GO:006448; P:regulation of translational elongation; IMP.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR PRODOM; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Repeat.
FT NP_BIND 232 239 ATP (POTENTIAL).
FT NP_BIND 565 572 ATP (POTENTIAL).
SQ SEQUENCE 752 AA; 85027 MW; 486FED10305A572E CRC64;

Query Match 42.1%; Score 40; DB 1; Length 752;
Best Local Similarity 37.5%; Pred. No. 63;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 RIAGKGRKLVYDSAR 16
DB 152 KVAKRNKFKVYKASK 167

RESULT 22
DYN1 CAEEL STANDARD; PRT; 830 AA.
AC P39055; Q95QY9;
DT 01-FEB-1995 (Rel. 31, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DYN1 OR C02C6.1.
GN DYN-1 OR C02C6.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=Bristol N2;
RX MEDLINE=97439883; PubMed=9284229;
RA Clark S.G.; Shurland D.L.; Meyerowitz E.M.; Bargmann C.I.;
RA van der Bliek A.M.;
RT "A dynamin GTPase mutation causes a rapid and reversible temperature-
RT inducible locomotion defect in C. elegans."
RN Proc. Natl. Acad. Sci. U.S.A. 94:10438-10443(1997).
RN [2]
RP REVISIONS TO C-TERMINUS.
RA van der Bliek A.M.;
RA Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: Microtubule-associated force-producing protein involved
CC in producing microtubule bundles and able to bind and hydrolyze
CC GTP. Most probably involved in vesicular trafficking processes, in
CC particular endocytosis.

```

```

CC -1- CATALYTIC ACTIVITY: GTP + H(2)O = GDP + phosphate.
CC -1- SUBCELLULAR LOCATION: Microtubule-associated.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=P39055-1; Sequence=Displayed;
CC Name=b;
CC IsoId=P39055-2; Sequence=VSP_001329;
CC Notes=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the dynamin family.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U29031; AAB72228.2; -
DR EMBL; Z79596; CAB01857.1; -
DR EMBL; Z79596; CAC42251.1; -
DR PIR; T18860; T18860.
DR HSSP; Q05193; 2DYN.
DR WormPep; C02C6.1a; CE07833.
DR WormPep; C02C6.1b; CE07832.
DR InterPro; IPR001401; Dynamin.
DR InterPro; IPR00375; Dynamin_central.
DR InterPro; IPR003130; GED.
DR InterPro; IPR001849; PH.
DR Pfam; PF00350; dynamin; 1.
DR Pfam; PF01031; dynamin_2; 1.
DR Pfam; PF02212; GED; 1.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00195; DYNAMIN.
DR SMART; SM00053; DYNG; 1.
DR SMART; SM00302; GED; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS00410; DYNAMIN; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW Hydrolyase; Motor protein; GTP-binding; Microtubule; Multigene family;
KW Endocytosis; Alternative splicing.
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 138 142 GTP (BY SIMILARITY).
FT NP_BIND 207 210 GTP (BY SIMILARITY).
FT DOMAIN 519 624 PH.
FT VARSPPLIC 817 830
FT VSPGPPPPNMAPPR -> VPTPSNGAPEIPARQVPRPF
FT (in isoform b).
FT /FTID=VSP_001329.
FT R -> P (IN REF. 1).
SQ SEQUENCE 830 AA; 93407 MW; PF681250E51AB8A5 CRC64;

Query Match 42.1%; Score 40; DB 1; Length 830;
Best Local Similarity 46.2%; Pred. No. 70;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 KEGRKLVDYDSAR 16
DB 88 KKGHRFVDFDAVR 100

RESULT 23
Y544 METJA STANDARD; PRT; 229 AA.
AC Q57964;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0544.
GN MJ0544.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;

```

CC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073 (1996).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67504; AAB98538.1; --  
 DR F1R; H64367; H64367.  
 DR TIGR; M00544; --  
 DR InterPro; IPR001173; Glyco trans 2; 1.  
 DR Pfam; PF00535; Glycos transf 2; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 229 AA; 26470 MW; D276082D92FCB26F CRC64;  
 Query Match 41.6%; Score 39.5; DB 1; Length 229;  
 Best Local Similarity 42.1%; Pred. No. 21;  
 Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;  
 QY 1 RIAGRGKLV-DYDSARHH 18  
 Db 185 QLAQKAGKIVRYVSEYH 203  
 RESULT 24  
 LEU3\_RALSO  
 ID LEU3\_RALSO STANDARD; PRT; 356 AA.  
 AC Q8XXF5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)  
 DE (IMDH) (3-IPM-DH)  
 GN LEUB OR LEUB1 OR RSC1988 OR RS03417.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=GMI1000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,  
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choine N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Molzan A., Robert C., Saurin W., Schiex T.,  
 RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weisenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
 RL Nature 415:497-502 (2002).  
 CC -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-  
 CC methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-  
 CC oxopentanoate. The product decarboxylates to 4-methyl-2-  
 CC oxopentanoate.

CC -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +  
 CC NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.  
 CC -!- PATHWAY: Leucine biosynthesis; third step.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate  
 CC dehydrogenase family. Leub subfamily 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL646067; CAD15690.1; ALT\_INIT.  
 DR HAMAP; MF 01033; -- 1.  
 DR InterPro; IPR001804; Isodh.  
 DR InterPro; IPR004429; Leub.  
 DR Pfam; PF00180; isodh; 1.  
 DR TIGR; TIGR00169; leub; 1.  
 DR PROSITE; PS00470; IDH IMDH; 1.  
 KW Oxidoreductase; Leucine biosynthesis; NAD; Complete proteome.  
 SQ SEQUENCE 356 AA; 38218 MW; 3FFA86BACD020E8A CRC64;  
 Query Match 41.1%; Score 39; DB 1; Length 356;  
 Best Local Similarity 61.5%; Pred. No. 41;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 3 AKRGKLVVDYDSA 15  
 Db 180 AKRGKLVVDYDKA 192  
 RESULT 25  
 MSL3\_DROME  
 ID MSL3\_DROME STANDARD; PRT; 512 AA.  
 AC P50336; Q81Q74; Q9VSI8;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Male-specific lethal-3 protein.  
 GN MSL-3 OR CG8631.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=95285767; PubMed=7768187;  
 RA Gorman M., Franke A., Baker B.S.;  
 RT "Molecular characterization of the male-specific lethal-3 gene and  
 RT investigations of the regulation of dosage compensation in  
 RT Drosophila.";  
 RL Development 121:463-475 (1995).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M.E., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mankulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "the genome sequence of *Drosophila melanogaster*.";  
 RA Science 287:2185-2195 (2000).  
 RA [3]  
 RA REVISIONS, AND ALTERNATIVE SPLICING.  
 RA MEDLINE=2426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RA systematic review";  
 RA Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RA [4]  
 RA SEQUENCE FROM N.A. (ISOFORM 1).  
 RA STRAIN=Berkeley; TISSUE=Embryo;  
 RA MEDLINE=2426066; PubMed=12537569;  
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Celniker S.E.;  
 RA "A *Drosophila* full-length cDNA resource";  
 RA Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 CC -1- FUNCTION: The MSL proteins are essential for elevating  
 CC transcription of the single X chromosome in the male (X chromosome  
 CC dosage compensation). Mle, msl-1 and msl-3 are co-localized on the  
 CC X chromosome. Each of the MSL proteins requires all the other MSLs  
 CC for wild-type X-chromosome binding.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; MSL-3 IS ASSOCIATED WITH HUNDREDS  
 CC OF DISCRETE SITES ALONG THE LENGTH OF THE X CHROMOSOME IN MALES  
 CC AND NOT IN FEMALES, AND IS ALSO ASSOCIATED WITH 10-20 AUTOSOMAL  
 CC SITES IN MALES.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=A;  
 CC IsoId=P50536-1; Sequences=Displayed;  
 CC Name=2; Synonyms=B;  
 CC IsoId=P50536-2; Sequences=VSP 007640;  
 CC Note=No experimental confirmation available;  
 CC -1- SIMILARITY: Contains 1 chromo domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC DR EMBL; X81321; CAAS7101.1; -  
 DR EMBL; AE003560; AAF50612.1; -  
 DR EMBL; AE003560; AAN2073.1; -  
 DR EMBL; AY071074; AAL48696.1; -  
 DR PIR; S48828; S48828.  
 DR FlyBase; FBgn0002775; msl-3.  
 DR GO; GO:0016456; C:dosage compensation complex (sensu *Drosophila*); NAS.  
 DR GO; GO:0003682; P:chromatin binding; IDA.  
 DR GO; GO:0009047; P:dosage compensation, by hyperactivation of . . . ; NAS.  
 DR InterPro; IPR000953; Chromo.  
 DR InterPro; IPR008676; MRG.  
 DR Pfam; PF05712; MRG; 1.  
 DR SMART; SM00298; CHROMO; 1.  
 DR PROSITE; PS00598; CHROMO 1; FALSE NEG.  
 DR PROSITE; PS50013; CHROMO 2; FALSE NEG.  
 KW Chromatin regulator; Nuclear protein; Transcription regulation;  
 KW Alternative splicing.  
 KW DOMAIN 11 90  
 FT VARSPLIC 1 34  
 FT CHROMO.  
 FT MTELRDETFLPHKGEIVLCYEPDKSKARVLYTSK -> MPQ  
 FT (in isoform 2).  
 FT /FTid=VSP 007640.  
 FT CONFLICT 184 184 A -> T (IN REF. 1).  
 FT CONFLICT 352 352 V -> G (IN REF. 1).  
 FT SEQUENCE 512 AA; 58852 MW; 3101BC92B2339951 CRC64;  
 Query Match 41.1%; Score 39; DB 1; Length 512;  
 Best Local Similarity 46.2%; Pred. No. 61;  
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RIATKGRKLVLDYD 13  
 DB 203 RVSERLRELYED 215  
 RESULT 26  
 ANRS HUMAN STANDARD; PRT; 776 AA.  
 ID Q3NU02; Q9H6Y9;  
 AC 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Ankyrin repeat domain protein 5.  
 GN ANKRD5.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark A.G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leharasliatho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RN Nature 414:865-871(2001).  
 RN  
 RP SEQUENCE OF 329-776 FROM N.A.  
 RC TISSUE=Colon;  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 8 ANK repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL109754; CAB87618.2;  
 DR EMBL; AK025322; BAB15111.1; ALT\_INIT.  
 DR HSP; P80144; 2MYO.  
 DR Genew; HGNC:15803; ANKRD5.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR002048; EF-hand.  
 DR Pfam; PF00023; ank; 10.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 11.  
 DR PROSITE; PS00888; ANK\_REPEAT; 6.  
 DR PROSITE; PS0297; ANK\_REPEAT\_REGION; 1.  
 KW ANK repeat; Repeat.  
 FT REPEAT 47 76 ANK 1.  
 FT REPEAT 184 213 ANK 2.  
 FT REPEAT 217 246 ANK 3.  
 FT REPEAT 250 279 ANK 4.  
 FT REPEAT 524 553 ANK 5.  
 FT REPEAT 557 586 ANK 6.  
 FT REPEAT 590 619 ANK 7.  
 FT REPEAT 623 652 ANK 8.  
 SQ SEQUENCE 776 AA; 86664 MW; 2F71F35AC4D337B6 CRC64;  
 Query Match 41.1%; Score 39; DB 1; Length 776;  
 Best Local Similarity 41.2%; Pred. No. 96;  
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 OY 2 IAKRGKLVYDSARHH 18  
 DB 204 ILERGEVNAFNDRRH 220  
 RESULT 27  
 ID QAL5 NEUCR STANDARD; PRT; 918 AA.  
 AC P11637;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DE Quinate repressor.  
 GN QA-1S.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RX MEDLINE=89293848; PubMed=2525625;  
 RA Geever R.F., Huie L., Baum J.A., Tyler B.M., Patel V.B.,  
 RA Rutledge B.J., Case M.E., Giles N.H.;

RT "DNA sequence, organization and regulation of the qa gene cluster of  
 RT Neurospora crassa.";  
 RL J. Mol. Biol. 207:15-34(1989).  
 CC -1- FUNCTION: Repressor for enzymes and proteins of quinate  
 CC metabolism.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X14603; CAA32753.1;  
 DR PIR; S04255; S04255.  
 DR InterPro; IPR001381; DHQuinase I.  
 DR InterPro; IPR000623; Shik\_kinase.  
 DR InterPro; IPR006151; Shikimate\_DH.  
 DR Pfam; PF01487; DHQuinase I; 1.  
 DR Pfam; PF01488; Shikimate\_DH; 1.  
 DR Pfam; PF01202; SKI; 1.  
 KW Quinate metabolism; Transcription regulation; Repressor; DNA-binding.  
 SQ SEQUENCE 918 AA; 100580 MW; 67EDA399CBF098B2 CRC64;  
 Query Match 41.1%; Score 39; DB 1; Length 918;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 OY 7 RKLVDYDSARHH 18  
 DB 127 RKIVLESEFHH 138  
 RESULT 28  
 ID POLG LIV STANDARD; PRT; 980 AA.  
 AC P22338;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Genome polyprotein [Contains: Capsid protein C (Core protein); Matrix  
 DE protein (Envelope protein M); Major envelope protein E; Nonstructural  
 DE proteins NS1] (Fragment).  
 OS Louping ill virus [Lil].  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus.  
 OX NCBI\_TaxID=11086;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91082437; PubMed=1845834;  
 RA Shiu S.Y.W., Ayres M.D., Gould E.A.;  
 RT "Genomic sequence of the structural proteins of louping ill virus:  
 RT comparative analysis with tick-borne encephalitis virus.";  
 RL Virology 180:411-415(1991).  
 CC -1- SUBUNIT: The virion of this virus is a nucleocapsid covered by a  
 CC lipoprotein envelope. The envelope consists of two proteins:  
 CC protein M and glycoprotein E. The nucleocapsid is a complex of  
 CC protein C and RNA.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M59376; AAA46281.1;  
 DR PIR; A38523; A38523.  
 DR HSP; P14336; 1SVB.  
 DR InterPro; IPR001122; Flavi\_capsidC.  
 DR InterPro; IPR000336; Flavi\_glycoproteE.



DR InterPro: IPR000069; Flavi M.  
 DR InterPro: IPR001157; Flavi NS1.  
 DR InterPro: IPR002535; Flavi Propep.  
 DR InterPro: IPR007110; IG-like.  
 DR Pfam: PF01003; Flavi Capsid; 1.  
 DR Pfam: PF02832; Flavi Glycop C; 1.  
 DR Pfam: PF00869; Flavi Glycoprot; 1.  
 DR Pfam: PF01004; Flavi M; 1.  
 DR Pfam: PF00948; Flavi NS1; 1.  
 DR Pfam: PF01570; Flavi Propep; 1.  
 DR ProDom: PD001556; Flavi Glycoprote; 1.  
 DR ProDom: PD001496; Flavi NS1; 1.  
 KW Polyprotein; Glycoprotein; Core protein; Coat protein;  
 Envelope protein; Transmembrane; Nonstructural protein.  
 FT INIT\_MET 1  
 FT CHAIN 1 112  
 FT PROPEP 113 205  
 FT CHAIN 206 280  
 FT CHAIN 281 776  
 FT CHAIN 777 >980  
 FT TRANSMEM 101 112  
 FT TRANSMEM 247 259  
 FT TRANSMEM 266 280  
 FT TRANSMEM 738 751  
 FT DISULFID 283 310  
 FT DISULFID 340 386  
 FT DISULFID 354 385  
 FT DISULFID 372 401  
 FT DISULFID 466 570  
 FT DISULFID 587 618  
 FT CARBOHYD 144 144  
 FT CARBOHYD 434 434  
 FT CARBOHYD 861 861  
 FT NON\_TER 980 980  
 SQ SEQUENCE 980 AA; 107234 MW; 2242E99364F3BFD9 CRC64;

Query Match 41.1%; Score 39; DB 1; Length 980;  
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 RGRKLVYDSARHH 18  
 |||| : : :  
 DB 220 RGRKWLGDLSLRTH 233

RESULT 29  
 YN48\_YEAST STANDARD; PRT; 981 AA.  
 AC Q12751;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical 113.2 kDa protein in S902-HSC82 intergenic region.  
 GN YNR185W OR YMR010.15.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=288c / AB572;  
 RX MEDLINE=97313268; PubMed=9169872;  
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
 RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,  
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 XIII.";  
 RL Nature 387:90-93 (1997).  
 RN [2]  
 RP SEQUENCE OF 650-981 FROM N.A.  
 RX MEDLINE=95373281; PubMed=7645348;

RA Erkine A.M., Szent-Gyorgyi C., Simmons S.F., Gross D.S.;  
 RT "The upstream sequences of the HSP82 and HSC82 genes of Saccharomyces  
 RT cerevisiae: regulatory elements and nucleosome positioning motifs.";  
 RL Yeast 11:573-580 (1995).  
 RN [3]  
 RP SEQUENCE OF 976-981 FROM N.A.  
 RX MEDLINE=89384620; PubMed=2674684;  
 RA Borkovich K.A., Farrelly F.W., Finkelstein D.B., Taulien J.,  
 RA Lindquist S.;  
 RT "hsp82 is an essential protein that is required in higher  
 RT concentrations for growth of cells at higher temperatures.";  
 RL Mol. Cell. Biol. 9:3919-3930 (1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z49808; CAA89918.1;  
 DR EMBL; U20349; AAA87904.1; ALT\_INIT.  
 DR PIR; S55132; S55132.  
 DR GerMOnline; 142859;  
 DR SGD; S0004797; YMR185W.  
 DR InterPro: IPR008938; ARM.  
 KW Hypothetical protein.  
 SQ SEQUENCE 981 AA; 113248 MW; 13F29F65C6D4E53B CRC64;

Query Match 41.1%; Score 39; DB 1; Length 981;  
 Best Local Similarity 40.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARHH 18  
 |||| : : :  
 DB 142 KKGOKSAEYETPRWH 156

RESULT 30  
 YD96\_METJA STANDARD; PRT; 2894 AA.  
 AC Q58791;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hypothetical protein MJ1396.  
 GN MJ1396.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8689087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.P., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073 (1996).  
 CC -----  
 CC -!- SIMILARITY: Contains 20 Pbl1 repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M93438; AAA34255.1; -  
 DR PIR; J01247; J01247.  
 DR HSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; Ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGR; TIGR00231; Small GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport;  
 KW Multigene family.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).  
 FT LIPID 202 202 S-geranylgeranyl cysteine  
 FT LIPID 203 203 (By similarity).  
 FT LIPID 203 203 S-geranylgeranyl cysteine  
 FT LIPID 203 203 (By similarity).  
 SQ SEQUENCE 203 AA; 22532 MW; CD143BADB92B4C65 CRC64;  
 Query Match 40.0%; Score 38; DB 1; Length 203;  
 Best Local Similarity 54.5%; Pred. No. 33;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 GRKLVYDSAR 16  
 :|:|:|:|:|:  
 Db 127 GKVDVYQAK 137

RESULT 32  
 ID FL1 EUCGL STANDARD; PRT; 359 AA.  
 AC 064953;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Leafy/floricaula homolog FL1 (ELF1).  
 GN LFI.  
 OS Eucalyptus globulus (Blue gum).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Myrtales; Myrtaceae; Eucalyptus.  
 OX NCBI\_TaxID=34317;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98363643; PubMed=9700063;  
 RA Southern S.G., Strauss S.H., Olive M.R., Harcourt R.L., Decroocq V.,  
 RA Zhu X., Llewellyn D.J., Peacock W.J., Dennis E.S.;  
 RT "Eucalyptus has a functional equivalent of the Arabidopsis floral  
 meristem identity gene LEAFY.";  
 RL Plant Mol. Biol. 37:897-910(1998).  
 CC -!- FUNCTION: Probable transcription factor (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- TISSUE SPECIFICITY: Expressed strongly in the early floral  
 CC primordium and then successively in the primordia of sepals,  
 CC petals, stamens and carpels. Also in the leaf primordia and young  
 CC leaves.  
 CC -!- SIMILARITY: Belongs to the FLO / LFY family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M93438; AAA34255.1; -  
 DR PIR; J01247; J01247.  
 DR HSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; Ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGR; TIGR00231; Small GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport;  
 KW Multigene family.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).  
 FT LIPID 202 202 S-geranylgeranyl cysteine  
 FT LIPID 203 203 (By similarity).  
 FT LIPID 203 203 S-geranylgeranyl cysteine  
 FT LIPID 203 203 (By similarity).  
 SQ SEQUENCE 203 AA; 22532 MW; CD143BADB92B4C65 CRC64;  
 Query Match 41.1%; Score 39; DB 1; Length 2894;  
 Best Local Similarity 61.5%; Pred. No. 4e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RIARGRKLVYD 13  
 :|:|:|:|:|:  
 Db 2068 KILNAGVKLVYD 2080

RESULT 31  
 ID YPT1 VOLCA STANDARD; PRT; 203 AA.  
 AC P31584;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE GTP-binding protein yptV1.  
 GN YPTV1.  
 OS Volvox carteri.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Volvocaceae; Volvox.  
 OX NCBI\_TaxID=3067;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=f. Nagariensis / HK10;  
 RX MEDLINE=92380499; PubMed=1511889;  
 RA Fabry S., Naeas N., Huber H., Palme K., Jaenicke L., Schmitt R.;  
 RT "The yptV1 gene encodes a small G-protein in the green alga Volvox  
 carteri: gene structure and properties of the gene product.";  
 RL Gene 118:153-162(1992).  
 CC -!- FUNCTION: Protein transport. Probably involved in vesicular  
 CC traffic (By similarity).  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its

```

CC -----
DR EMBL: AF034806; AAC31359.1; --
DR InterPro: IPR002910; FLO_LFY.
DR Pfam: PF01698; FLO_LFY; 1.
KW Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; Developmental protein.
SQ SEQUENCE 359 AA; 40929 MW; C9C768BE85076312 CRC64;

Query Match 40.0%; Score 38; DB 1; Length 359;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 RKLVDYDSARH 18
   ||:| |
Db 84 RLLLEADRRH 95

RESULT 33
ARGU_BRAJA STANDARD; PRT; 413 AA.
AC PS9610;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Arginine biosynthesis bifunctional protein argJ [Includes: Glutamate
DE N-acetyltransferase (EC 2.3.1.35) (Ornithine acetyltransferase)
DE (ornithine transacetylase) (OATase); Amino-acid acetyltransferase
DE (EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)] [Contains: Arginine
DE biosynthesis bifunctional protein argJ alpha chain; Arginine
DE biosynthesis bifunctional protein argJ beta chain].
OS ARGJ OR BLA0206.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RA "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -!- FUNCTION: Catalyzes two activities which are involved in the
CC cyclic version of arginine biosynthesis: the synthesis of
CC acetylglutamate from glutamate and acetyl-CoA, and of ornithine by
CC transacetylation between acetylornithine and glutamate (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-
CC ornithine + N-acetyl-L-glutamate.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-
CC glutamate.
CC -!- PATHWAY: Arginine biosynthesis; first step.
CC -!- PATHWAY: Arginine biosynthesis; fifth step.
CC -!- SUBUNIT: Heterotetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,
CC capable of catalyzing only the fifth step of the arginine
CC biosynthetic pathway.
CC -!- SIMILARITY: Belongs to the argJ family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

DR EMBL: AP005935; BAC45471.1; --
DR HAMAP; MF_01106; --; 1.
DR InterPro: IPR002813; ArgJ.
DR Pfam: PF01960; ArgJ; 1.
DR ProDom: PD004193; ArgJ; 1.
DR TIGRFAMs: TIGR00120; ArgJ; 1.
KW Arginine biosynthesis; Multifunctional enzyme; Transferase;
KW Acyltransferase; Complete proteome.
FT CHAIN 1 194
   ARGinine BIOSYNTHESIS BIFUNCTIONAL
   PROTEIN ARGJ ALPHA CHAIN (BY SIMILARITY).
FT CHAIN 195 413
   ARGinine BIOSYNTHESIS BIFUNCTIONAL
   PROTEIN ARGJ BETA CHAIN (BY SIMILARITY).
FT SITE 194 195
   CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).
SQ SEQUENCE 413 AA; 42683 MW; F5B3BE3F558DCF33 CRC64;

Query Match 40.0%; Score 38; DB 1; Length 413;
Best Local Similarity 43.8%; Pred. No. 72;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RIAKGRKRLVDYDSAR 16
   ||:| |
Db 355 RVARGARDPDYDEAQ 370

RESULT 34
PDA6 RAT STANDARD; PRT; 431 AA.
AC Q63081;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein disulfide isomerase A6 precursor (EC 5.3.4.1) (Protein
DE disulfide isomerase P5) (Calcium-binding protein 1) (CaBP1
DE (Fragment)).
OS PDI A6 OR CABP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=95181574; PubMed=7876340;
RA Fuelekruug J., Soennichsen B., Wuensch U., Arseven K., Van P.N.,
RA Soeling H.-D., Mieskes G.;
RA "CaBP1, a calcium binding protein of the thioredoxin family, is a
RT resident KDEL protein of the ER and not of the intermediate
RT compartment.";
RL J. Cell Sci. 107:2719-2727(1994).
CC -!- CATALYTIC ACTIVITY: Rearrangement of both intrachain and
CC interchain disulfide bonds in proteins to form the native
CC structures.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -!- SIMILARITY: Contains 2 thioredoxin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

DR PROSITE; PS00014; ER TARGET; 1.  
 DR PROSITE; PS00194; THIOREDOXIN; 2.  
 KW Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal;  
 KW Calcium-binding.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 10  
 FT CHAIN 11 431  
 FT DISULFID 46 49  
 FT DISULFID 181 184  
 FT DOMAIN 413 427  
 FT SITE 428 431  
 FT SEQUENCE 431 AA; 47320 MW; 2CD07A233C90135C CRC64;  
 Query Match 40.0%; Score 38; DB 1; Length 431;  
 Best Local Similarity 43.8%; Pred. No. 75;  
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 RIAGKGRKLVYDGSAR 16  
 DB 232 KIFQGESPVYDGGK 247  
 RESULT 35  
 PDA6 HUMAN  
 ID PDA6 HUMAN STANDARD; PRT; 440 AA.  
 AC Q15084; Q99778;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein disulfide isomerase A6 precursor (EC 5.3.4.1) (Protein  
 DE disulfide isomerase P5).  
 GN PDA6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=96069616; PubMed=7590364;  
 RA Hayano T., Kikuchi M.;  
 RT "Cloning and sequencing of the cDNA encoding human P5";  
 RL Gene 164:377-378 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny K.C., Harte S., Kettman E.J., Lu X., Gibbs R.A.,  
 RA Faney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP SEQUENCE OF 20-440 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97264341; PubMed=9110174;  
 RA Yu W., Anderson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,

RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;  
 RT "Large-scale concatenation cDNA sequencing.";  
 RL Genome Res. 7:353-358 (1997).  
 CC -1- CATALYTIC ACTIVITY: Rearrangement of both intrachain and  
 CC interchain disulfide bonds in proteins to form the native  
 CC structures.  
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).  
 CC -1- SIMILARITY: Contains 2 thioredoxin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; D49489; BAA08450.1; -;  
 DR EMBL; BC001312; AAH01312.1; -;  
 DR EMBL; U79278; AAB50217.1; -;  
 DR PIR; JC4369; JC4369.  
 DR HSP; P07237; IMEX.  
 DR GO; GO:0003756; P:protein disulfide isomerase activity; TAS.  
 DR GO; GO:0006457; P:protein folding; TAS.  
 DR InterPro; IPR005788; Disulph\_isom.  
 DR InterPro; IPR000886; ER\_target\_S.  
 DR InterPro; IPR006662; Thioered.  
 DR InterPro; IPR006663; Thioeredox\_dom2.  
 DR Pfam; PF00085; Thioered; 2.  
 DR PRINTS; PR00421; THIOREDOXIN.  
 DR TIGRFAMs; TIGR01126; pdi\_dom; 2.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 DR PROSITE; PS00194; THIOREDOXIN; 2.  
 DR Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 440 PROTEIN DISULFIDE ISOMERASE A6.  
 FT DISULFID 55 58 REDOX-ACTIVE (BY SIMILARITY).  
 FT DISULFID 190 193 REDOX-ACTIVE (BY SIMILARITY).  
 FT DOMAIN 422 434 ASP/GLU-RICH (ACIDIC).  
 FT SITE 437 440 PREVENT SECRETION FROM ER (POTENTIAL).  
 FT CONFLICT 214 214 K -> R (IN REF. 3).  
 SQ SEQUENCE 440 AA; 48121 MW; 06895409F0265D7C CRC64;  
 Query Match 40.0%; Score 38; DB 1; Length 440;  
 Best Local Similarity 43.8%; Pred. No. 77;  
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 RIAGKGRKLVYDGSAR 16  
 DB 241 KIFQGESPVYDGGK 256  
 RESULT 36  
 WCA6 ECOLI  
 ID WCA6 ECOLI STANDARD; PRT; 464 AA.  
 AC P71244; P76378;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Colanic acid biosynthesis protein wcaM.  
 GN WCA6 OR B2043.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=96326333; PubMed=8759852;  
 RA Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;  
 RT "Organization of the Escherichia coli K-12 gene cluster responsible  
 RT for production of the extracellular polysaccharide colanic acid";  
 RL J. Bacteriol. 178:4885-4893 (1996).

RN [2]  
 RP REVISIONS.  
 RC STRAIN=K12;  
 RA Reeves P.R.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12";  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,  
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,  
 RA Saito N., Sampei G., Seki Y., Sivasubram S., Tgami H.,  
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;  
 RT "A 460-Kb DNA sequence of the *Escherichia coli* K-12 genome  
 corresponding to the 40.1-50.0 min region on the linkage map";  
 RN [5]  
 RP PATHWAY: Slime polysaccharide colanic acid biosynthesis.  
 CC -  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U38473; AAC77852.1; -;  
 DR EMBL; AE000295; AAC75104.1; -;  
 DR EMBL; D90842; BAA15897.1; -;  
 DR PIR; B64970; B64970.  
 DR EcoGene; EGI2651; wcaM.  
 KW Lipopolysaccharide biosynthesis; Complete proteome.  
 SQ SEQUENCE 464 AA; 51315 MW; 72A7655DC07368BE CRC64;  
 Query Match 40.0%; Score 38; DB 1; Length 464;  
 Best Local Similarity 35.3%; Pred. No. 82;  
 Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
 QY 2 IAKRGKLYDYSARHH 18  
 DB 432 VNEKGSVDIDRINH 448  
 : : : : :  
 RESULT 37  
 YH33 CHRV  
 ID YH33 CHRV STANDARD; PRT; 478 AA.  
 AC QYNX94;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hypothetical UPF0061 protein CV1733.  
 GN CV1733.  
 OS Chromobacterium violaceum.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Chromobacterium.  
 OX NCBI\_TaxID=536;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12472 / DSM 30191;  
 RX MEDLINE=2288280; PubMed=14500782;  
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,

RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,  
 RA Alves-Gomes J.A., Andrade E.M., Azevedo J., Baptista A.J., Bataus L.A.M.,  
 RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,  
 RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Buriti H.A.,  
 RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,  
 RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,  
 RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,  
 RA Fantinatti F., Farias L.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,  
 RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,  
 RA Gazinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,  
 RA Grattapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,  
 RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,  
 RA Madeira H.M.F., Manfio G.P., Maranhao A.O., Martins W.S.,  
 RA di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,  
 RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,  
 RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,  
 RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,  
 RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,  
 RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunanez H.N.,  
 RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,  
 RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,  
 RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,  
 RA Vettore A., Wassem K., Zaha A., Simpson A.J.G.;  
 RT "The complete genome sequence of *Chromobacterium violaceum* reveals  
 RT remarkable and exploitable bacterial adaptability";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).  
 CC -  
 CC -  
 CC -  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AE016916; AA059408.1; -;  
 DR HAMAP; MF\_00692; -; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 478 AA; 52250 MW; 34990A21225CED51 CRC64;  
 Query Match 40.0%; Score 38; DB 1; Length 478;  
 Best Local Similarity 63.6%; Pred. No. 84; Mismatches 1; Indels 0; Gaps 0;  
 Matches 7; Conservative 1;  
 QY 8 KLVYDYSARHH 18  
 DB 185 KLVYDYSARHH 195  
 : : : : :  
 RESULT 38  
 YH18 TREPA  
 ID YH18 TREPA STANDARD; PRT; 504 AA.  
 AC O83981;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical UPF0144 protein TP1018.  
 GN TP1018.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nichols;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
 RA McDonald L., Arliach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,

```

RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RL spirochete.";
RL Science 281:375-388(1998).
CC -!- SIMILARITY: Belongs to the UPF0144 family.
CC -!- SIMILARITY: Contains 1 HD domain.
CC -!- SIMILARITY: Contains 1 KH domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001269; AAC65967.1; -.
CC FIR; F71253; F71253.
CC TIGR; TP1018; -.
CC HAMAP; MF_00335; -.
CC InterPro; IPR006674; HD.
CC InterPro; IPR004087; KH dom.
CC InterPro; IPR004088; KH_type_1.
CC InterPro; IPR003607; Mac_puhsphidro.
CC InterPro; IPR006675; Unchar_HDIG.
CC Pfam; PF01966; KH; 1.
CC Pfam; PF00013; KH; 1.
CC SMART; SM00471; HDC; 1.
CC SMART; SM00322; KH; 1.
CC TIGRFAMs; TIGR00277; HDIG; 1.
CC PROSITE; PS50084; KH_TYPE_1; 1.
KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.
FT DOMAIN 2 22 POTENTIAL.
FT DOMAIN 193 261 KH.
FT DOMAIN 320 413 HD.
SQ SEQUENCE 504 AA; 57070 MW; D61EEE69B6434745 CRC64;

Query Match 40.0%; Score 38; DB 1; Length 504;
Best Local Similarity 47.8%; Pred. No. 89;
Matches 11; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

QY 2 IAKGRKLVYDVSARHH 18
| | | | |
DB 344 IAKGALLHDVIGKGVETSDRNH 366

RESULT 39
PHR NEUCR
ID PHR NEUCR STANDARD; PRT; 642 AA.
AC P27526;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Deoxyribodipyrimidine photolyase (EC 4.1.99.3) (DNA photolyase)
DE (Photoreactivating enzyme).
DE PHR.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020228; PubMed=1833725;
RA Yajima H., Inoue H., Oikawa A., Yasui A.;
RT "Cloning and functional characterization of a eucaryotic DNA
RT photolyase gene from Neurospora crassa.";
RL Nucleic Acids Res. 19:5359-5362(1991).
CC -!- FUNCTION: This enzyme catalyzes the light-dependent monomerization
CC (300-600 nm) of cyclobutyl pyrimidine dimers (in cis-syn
CC configuration) which are formed between adjacent bases on the
CC same DNA strand, upon exposure to ultraviolet radiation.
CC -!- CATALYTIC ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine
CC residues (in DNA).

```

```

CC -!- COPACTOR: Contains 2 chromophores: a reduced flavin (FADH2) and a
CC 5,10-methylenetetrahydrofolate. Both chromophores are bound by non-
CC covalent interactions.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to the DNA photolyase class-1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X58713; CAA41549.1; -.
CC FIR; S18667; S18667.
CC DR HSHP; P00914; LDNP.
CC InterPro; IPR002081; DNA photolyase 1.
CC InterPro; IPR006050; DNA photolyase_N.
CC InterPro; IPR005101; FAD_binding_7.
CC InterPro; IPR006051; FAD_binding_N.
CC Pfam; PF00875; DNA photolyase; 1.
CC Pfam; PF03441; FAD_binding_7; 1.
CC PRINTS; P00147; DNAPHOTLYASE.
CC PRODOM; PD004390; FAD binding N; 1.
CC PROSITE; PS00394; DNA_PHOTOLYASES_1_1; 1.
CC PROSITE; PS00691; DNA_PHOTOLYASES_1_2; 1.
KW Lyase; Chromophore; Flavoprotein; FAD; DNA repair; DNA-binding.
SQ SEQUENCE 642 AA; 73076 MW; 05B1A56B69F77EDC CRC64;

Query Match 40.0%; Score 38; DB 1; Length 642;
Best Local Similarity 43.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 AKGRKLVYDVSARHH 18
| | | | |
DB 339 APEGKRLRDEKARYH 354

RESULT 40
NADO THEER
ID NADO THEER STANDARD; PRT; 651 AA.
AC P32382;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH oxidase (EC 1.-.-.-)
OS Thermoanaerobacter brockii (Thermoanaerobium brockii).
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=29323;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=RT8.G4;
RC MEDLINE=93363637; PubMed=8357835;
RA Liu X.-L., Scopes R.K.;
RT "Cloning, sequencing and expression of the gene encoding NADH oxidase
RT from the extreme anaerobic thermophile Thermoanaerobium brockii.";
RL Biochim. Biophys. Acta 1174:187-190(1993).
CC -!- FUNCTION: Reduces a range of alternative electron acceptors.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- COPACTOR: FAD; contains 2 FAD molecules per subunit.
CC -!- SUBUNIT: Homohexamer.
CC -!- PTM: The N-terminus is blocked.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```

```
DR EMBL: X67220; CAA47660.1; -.
DR F1R; S35706; S35706.
DR InterPro; IPR000759; Admrx_reductase.
DR InterPro; IPR001327; FAD pyr_redox.
DR InterPro; IPR003009; FWN enzyme.
DR InterPro; IPR000205; NAD BS.
DR InterPro; IPR001155; Oxidored_FMN.
DR InterPro; IPR001100; Pyr redox.
DR InterPro; IPR000103; Pyridine_redox_2.
DR InterPro; IPR000594; Thif_domain.
DR Pfam; PF00724; oxidored_FMN; 1.
DR Pfam; PF00070; Pyr redox; 1.
DR PRINTS; PRO0419; ADXRDASE.
DR PRINTS; PRO0368; FADPNR.
DR PRINTS; PRO0411; PNDRTASEI.
DR PRINTS; PRO0469; PNDRTASEII.
DR Oxidoreductase; NAD; FAD; Flavoprotein; Iron-sulfur; 4Fe-4S.
FT METAL 344 344
FT METAL 347 347
FT METAL 351 351
FT METAL 364 364
FT NP_BIND 387 413
FT NP_BIND 515 542
SQ SEQUENCE 651 AA; 71306 MW; F2282A24DA817439 CRC64;
Query Match 40.0%; Score 38; DB 1; Length 651;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 3 AKRGKLVYDVSARH 17
DB 405 AKRGHVILYKQKH 419
RESULT 41
STC2 STAAU STANDARD; PRT; 715 AA.
AC F17855;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Staphylocoagulase precursor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BB;
RX MEDLINE=88139269; PubMed=3481366;
RA Kaiga S., Miyata T., Yoshizawa Y., Kawabata S., Morita T.,
RA Igashiri H., Iwanaga S.;
RT "Nucleotide sequence of the staphylocoagulase gene: its unique COOH-terminal 8 tandem repeats."
RL J. Biochem. 102:1177-1186(1987).
CC -!- FUNCTION: STAPHYLOCOAGULASE IS AN EXTRACELLULAR PROTEIN WHICH SPECIFICALLY FORMS A COMPLEX WITH HUMAN PROTHROMBIN. THIS COMPLEX NAMED STAPHYLOTHROMBIN CAN CLOT FIBRINOGEN WITHOUT ANY PROTEOLYTIC CLEAVAGE OF PROTHROMBIN.
CC -!- DOMAIN: THE C-TERMINAL TANDEM REPEATS ARE NOT REQUIRED FOR THE PROCOAGULANT ACTIVITY.
CC -!- SIMILARITY: TO THE STAPHYLOCOAGULASE FROM STRAIN 213.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: D00184; BAA00126.1; -.
DR F1R; A41511; A41511.
DR InterPro; IPR001443; Staphylocoagulase.
```

```
DR Pfam; PF04022; Staphylocoagulase; 8.
DR PROSITE; PS00429; STAPHYLOCOAGULASE; 8.
KW Prothrombin activator; Repeat; Signal.
FT SIGNAL 1 26
FT CHAIN 27 715 STAPHYLOCOAGULASE.
FT DOMAIN 495 710 8 X 27 AA TANDEM REPEATS.
SQ SEQUENCE 715 AA; 80100 MW; 46ABC9567AF5F128 CRC64;
Query Match 40.0%; Score 38; DB 1; Length 715;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 9 LVDYDSARHH 18
DB 264 ITSVDSSKHH 273
RESULT 42
SECA MYCSM
ID SECA MYCSM STANDARD; PRT; 957 AA.
AC P71533;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Preprotein translocase seCA subunit.
GN SECA.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown A.B., Jacobs W.R.;
RT "Conservation of the general secretory pathway: isolation and characterization of SecA homologues from Mycobacterium bovis BCG and M. smegmatis."
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in protein export. Interacts with the secY/secE subunits. SecA has a central role in coupling the hydrolysis of ATP to the transfer of pre-secretory periplasmic and outer membrane proteins across the membrane (By similarity).
CC -!- SUBUNIT: Part of the prokaryotic protein translocation apparatus which comprise secA, secB, secD, secE, secF, secG and secY (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic side of membrane (By similarity).
CC -!- SIMILARITY: Belongs to the seCA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: U66081; AAB06754.1; -.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000185; SecA.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF01043; SecA protein; 1.
DR PRINTS; PRO0906; SECA.
DR TIGRFAMs; TIGR00963; seca; 1.
DR PROSITE; PS01312; SECA; 1.
KW Protein transport; ATP-binding; Membrane; Translocation; Transport.
FT NP_BIND 98 105 ATP (POTENTIAL).
SQ SEQUENCE 957 AA; 107001 MW; AD208569A22BA32E CRC64;
Query Match 40.0%; Score 38; DB 1; Length 957;
Best Local Similarity 53.3%; Pred. No. 1.8e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 RIAGRKGLVDYDSA 15
```





```

DR EMBL; U66362; -; NOT_ANNOTATED_CDS.
DR PIR; G01946; G01946.
DR HSP; P29476; 1QAV.
DR Genew; HGNC:7872; NOS1.
DR MIM; 163731.
DR GO; GO:0005624; Cmembrane fraction; TAS.
DR GO; GO:0004517; Finitric-oxide synthase activity; TAS.
DR GO; GO:0006936; Pmuscle contraction; TAS.
DR InterPro; IPR003097; FAD binding.
DR InterPro; IPR008254; Flav nitox synth.
DR InterPro; IPR001094; Flavodoxin like.
DR InterPro; IPR001709; FPN cyt reductase.
DR InterPro; IPR004030; NO synthase.
DR InterPro; IPR001433; Oxred_FAD/NAD(P).
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00667; FAD binding_1; 1.
DR Pfam; PF00258; flavodoxin; 1.
DR Pfam; PF00175; NAD binding_1; 1.
DR Pfam; PF02898; NO synthase; 1.
DR Pfam; PF00595; PDZ; 1.
DR PRINTS; PR00369; FLAVODOXIN.
DR PRINTS; PR00371; PFNCR.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS0902; FLAVODOXIN_LIKE; 1.
DR PROSITE; PS60001; NOS; 1.
DR PROSITE; PS50106; PDZ; 1.
DR Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme;
KW Alternative splicing; Multigene family.
FT DOMAIN 17 99
FT METAL 163 245
FT DOMAIN 163 245
FT METAL 760 940
FT DOMAIN 760 940
FT METAL 720 940
FT DOMAIN 720 940
FT NP_BIND 886 917
FT NP_BIND 1032 1185
FT NP_BIND 1175 1195
FT NP_BIND 1250 1268
FT NP_BIND 1348 1363
FT VARSPPLIC 1 336
FT VARSPPLIC 285 407
FT VARSPPLIC 408 1434
FT VARSPPLIC 509 613
FT CONFLICT 131 131
FT CONFLICT 178 184
FT CONFLICT 492 493
FT CONFLICT 549 549
FT CONFLICT 563 563
FT CONFLICT 1407 1407
FT CONFLICT 1434 AA; 160970 MW; 99235793B953BF37 CRC64;
SQ SEQUENCE 1434 AA; 160970 MW; 99235793B953BF37 CRC64;

Query Match 40.0%; Score 38; DB 1; Length 1434;
Best Local Similarity 62.5%; Pred. No. 2.8e-02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 2 IAKRGKLVLD--YDSA 15
Db 65 LAVNGEPLVDLSYDSA 80
: ||| ||| |||
: ||| ||| |||

RESULT 44
NOS1_RABIT
ID NOS1_RABIT
AC O19132;

```

```

DT DT 30-MAY-2000 (Rel. 39, Created)
DT DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DE Nitric-oxide synthase, brain (EC 1.14.13.39) (NOS, type I) (Neuronal
DE DE NOS) (N-NOS) (nNOS) (Constitutive NOS) (NC-NOS) (bNOS).
GN GN NOS1.
OS OS Oryctolagus cuniculus (Rabbit).
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX OX NCBI_TaxID=9986;
RN RN [1]
RP RP SEQUENCE FROM N.A.
RC RC TISSUE=Brain;
RA RA Jeong Y., Yim J.;
RT RT "Molecular cloning of a cDNA encoding a constitutive nitric oxide
RT RT synthase from rabbit brain.";
RL RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC CC -!- FUNCTION: Produces nitric oxide (NO) which is a messenger molecule
CC CC with diverse functions throughout the body. In the brain and
CC CC peripheral nervous system, NO displays many properties of a
CC CC neurotransmitter.
CC CC -!- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) = citrulline +
CC CC nitric oxide + N NADP(+).
CC CC -!- COFACTOR: Heme. Binds one mole each of FAD and FMN. Also requires
CC CC tetrahydrobiopterin (BH4) which may stabilize the dimeric form of
CC CC the enzyme (By similarity).
CC CC -!- ENZYME REGULATION: Stimulated by calcium/calmodulin. Inhibited by
CC CC n-NOS-inhibiting protein (PIN) which may prevent the dimerization
CC CC of the protein (By similarity).
CC CC -!- SUBUNIT: Homodimer. Forms a ternary complex with CAPON and RASD1.
CC CC -!- INTERACTIONS: Interacts with DLGA (By similarity).
CC CC -!- SUBCELLULAR LOCATION: In skeletal muscle, it is localized beneath
CC CC the sarcolemma of fast-twitch muscle fiber by associating with the
CC CC dystrophin glycoprotein complex (By similarity).
CC CC -!- DOMAIN: The PDZ domain in the N-terminal part of the neuronal
CC CC isoform participates in protein-protein interaction, and is
CC CC responsible for targeting nNOS to synaptic membranes in muscles
CC CC (By similarity).
CC CC -!- SIMILARITY: Belongs to the NOS family.
CC CC -!- SIMILARITY: Contains 1 flavodoxin-like domain.
CC CC -!- SIMILARITY: Contains 1 PDZ/DRH domain.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC CC or send an email to license@sib-sib.ch).
CC CC -----
CC CC EMBL; U91584; AB568663.1; -.
CC CC HSP; P29476; 1QAV.
CC CC InterPro; IPR003097; FAD binding.
CC CC InterPro; IPR008254; Flav nitox synth.
CC CC InterPro; IPR001094; Flavodoxin like.
CC CC InterPro; IPR001709; FPN cyt reductase.
CC CC InterPro; IPR004030; NO synthase.
CC CC InterPro; IPR001433; Oxred_FAD/NAD(P).
CC CC InterPro; IPR001478; PDZ.
CC CC Pfam; PF00667; FAD binding_1; 1.
CC CC Pfam; PF00258; flavodoxin; 1.
CC CC Pfam; PF00175; NAD binding_1; 1.
CC CC Pfam; PF02898; NO synthase; 1.
CC CC Pfam; PF00595; PDZ; 1.
CC CC PRINTS; PR00369; FLAVODOXIN.
CC CC PRINTS; PR00371; PFNCR.
CC CC SMART; SM00228; PDZ; 1.
CC CC PROSITE; PS0902; FLAVODOXIN_LIKE; 1.
CC CC PROSITE; PS60001; NOS; 1.
CC CC PROSITE; PS50106; PDZ; 1.
CC CC Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme;
KW Multigene family.
FT DOMAIN 17 99
FT VARSPPLIC 1 336
FT VARSPPLIC 285 407
FT VARSPPLIC 408 1434
FT VARSPPLIC 509 613
FT CONFLICT 131 131
FT CONFLICT 178 184
FT CONFLICT 492 493
FT CONFLICT 549 549
FT CONFLICT 563 563
FT CONFLICT 1407 1407
FT CONFLICT 1434 AA; 160970 MW; 99235793B953BF37 CRC64;
SQ SEQUENCE 1434 AA; 160970 MW; 99235793B953BF37 CRC64;

```

```

FT DOMAIN 164 246 NNOS-INHIBITING PROTEIN (PIN)-BINDING (BY
FT SIMILARITY).
FT FT 761 941 FLAVODOMAIN-LIKE.
FT FT 421 421 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT FT 731 751 CALMODULIN-BINDING (POTENTIAL).
FT FT 887 918 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 1033 1044 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 1176 1186 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 1251 1269 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 1349 1364 NADP (ADP PART) (BY SIMILARITY).
SQ SEQUENCE 1435 AA; 160864 MW; 3ED87ECDD93A7A5A CRC64;

Query Match 40.0%; Score 38; DB 1; Length 1435;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 2 IAKGRKLVLD--YDSA 15
DB 65 LAVNGRPLVDLSYDSA 80

RESULT 45
KKK1_YEAST
AC P34244; STANDARD; PRT; 1518 AA.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=94078677; PubMed=8256524;
RA Fallier C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
RA Bolutin-Fukuhara M.;
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT physically localizes the MRB1 gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RT kinases";
RL Yeast 9:1149-1155(1993).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X71133; CAA50456.1; -
CC DR EMBL; Z28101; CAA81941.1; -
CC DR EIR; S37928; S37928.
CC DR HSSP; Q63450; 1A06.
CC DR GerMOnline; 139857; -.
CC DR SGD; S0001584; HSL1.
CC DR GO; GO:0005935; C:bud neck; IDA.
CC DR GO; GO:0005940; C:septin ring; IDA.
CC DR GO; GO:0004672; P:protein kinase activity; IDA.
CC DR GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
CC DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
CC DR GO; GO:0000074; P:regulation of cell cycle; IMP.
CC DR GO; GO:0000135; P:septin checkpoint; IGI.
CC DR InterPro; IPR000719; Prot kinase.
CC DR InterPro; IPR008271; Ser_thr_kinase.
CC DR InterPro; IPR002290; Ser_thr_kinase.
CC DR InterPro; IPR001245; Tyr_kinase.

```

```

DR PFAM; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 81 369 PROTEIN_KINASE.
FT NP_BIND 87 95 ATP (BY SIMILARITY).
FT BINDING 110 110 ATP (BY SIMILARITY).
FT ACT_SITE 239 239 BY SIMILARITY.
SQ SEQUENCE 1518 AA; 169592 MW; 803F84F7531241DD CRC64;

Query Match 40.0%; Score 38; DB 1; Length 1518;
Best Local Similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 KRGRKLVLDYD 13
DB 1451 KFGKRVVEYD 1460

RESULT 46
SPCA_HUMAN
ID SPCA_HUMAN STANDARD; PRT; 2418 AA.
AC P02549; Q15514;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Spectrin alpha chain, erythrocyte (Srythroid alpha-spectrin).
GN SPTA1 OR SPTA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90170949; PubMed=1689726;
RA Sahr K.E., Laurila P., Kotula L., Scarpa A.L., Coupal E., Leto T.L.,
RA Linnenbach A.J., Winkelmann J.C., Speicher D.W., Marchesi V.T.,
RA Curtis P.J., Forget B.G.;
RT "The complete cDNA and polypeptide sequences of human erythroid
RT alpha-spectrin";
RL J. Biol. Chem. 265:4434-4443(1990).
RN [2]
RP SEQUENCE OF 7-533 FROM N.A., AND VARIANTS EL2 PRO-260; PRO-261 AND
RP PRO-471.
RX MEDLINE=90009318; PubMed=2794061;
RA Sahr K.E., Tobe T., Scarpa A., Laughinghouse K., Marchesi S.L.,
RA Agre P., Linnenbach A.J., Marchesi V.T., Forget B.G.;
RT "Sequence and exon-intron organization of the DNA encoding the alpha
RT I domain of human spectrin. Application to the study of mutations
RT causing hereditary elliptocytosis";
RL J. Clin. Invest. 84:1243-1252(1989).
RN [3]
RP SEQUENCE OF 7-601.
RX MEDLINE=84087888; PubMed=6654896;
RA Speicher D.W., Davis G., Marchesi V.T.;
RT "Structure of human erythrocyte spectrin. II. The sequence of the
RT alpha-I domain";
RL J. Biol. Chem. 258:14938-14947(1983).
RN [4]
RP SEQUENCE OF 7-125.
RX MEDLINE=84087887; PubMed=6654895;
RA Speicher D.W., Davis G., Yurchenco P.D., Marchesi V.T.;
RT "Structure of human erythrocyte spectrin. I. Isolation of the alpha-I
RT domain and its cyanogen bromide peptides";
RL J. Biol. Chem. 258:14931-14937(1983).
RN [5]
RP SEQUENCE OF 320-450 FROM N.A.
RX MEDLINE=86205962; PubMed=3458204;

```

RA Linnenbach A.J., Speicher D.W., Marchesi V.T., Forget B.G.;  
 RT "Cloning of a portion of the chromosomal gene for human erythrocyte  
 RL alpha-spectrin by using a synthetic gene fragment."  
 RN Proc. Natl. Acad. Sci. U.S.A. 83:2397-2401(1986).  
 [6]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=84295638; PubMed=6472478;  
 RA Speicher D.W., Marchesi V.T.;  
 RT "Erythrocyte spectrin is comprised of many homologous triple helical  
 RL segments."  
 RN Nature 311:177-180(1984).  
 [7]  
 RP IDENTIFICATION OF PROBABLE FRAMESHIFT IN 2407-2418.  
 RX MEDLINE=97001215; PubMed=8844207;  
 RA Maillet P., Alloisio N., Morle L., Delaunay J.;  
 RT "Spectrin mutations in hereditary elliptocytosis and hereditary  
 RL spherocytosis."  
 RN Hum. Mutat. 8:97-107(1996).  
 [9]  
 RP VARIANT EL2 SER-24.  
 RX MEDLINE=94289716; PubMed=8018926;  
 RA Parquet N., Devaux I., Boulanger L., Galand C., Boivin P.,  
 RT Lecomte M.-C., Dhery D., Garbarz M.;  
 RL "Identification of three novel spectrin alpha I/74 mutations in  
 RT hereditary elliptocytosis: further support for a triple-stranded  
 RL folding unit model of the spectrin heterodimer contact site."  
 RN Blood 84:303-308(1994).  
 [10]  
 RP VARIANTS EL2 CYS-28; HIS-28; LEU-28 AND SER-28.  
 RX MEDLINE=91358728; PubMed=1679439;  
 RA Coetzer T.L., Sahr K., Pechal J., Blacklock H., Peterson L., Koler R.,  
 RT Doyle J., Manaster J., Palek J.;  
 RL "Four different mutations in codon 28 of alpha spectrin are  
 RT associated with structurally and functionally abnormal spectrin alpha  
 RL I/74 in hereditary elliptocytosis."  
 RN J. Clin. Invest. 88:743-749(1991).  
 [11]  
 RP VARIANT EL2 SER-28, AND VARIANT HPP ARG-48.  
 RX MEDLINE=91346849; PubMed=1879597;  
 RA Floyd P.B., Gallagher P.G., Valentino L.A., Davis M., Marchesi S.L.,  
 RT Forget B.G.;  
 RL "Heterogeneity of the molecular basis of hereditary  
 RT pyropoikilocytosis and hereditary elliptocytosis associated with  
 RT increased levels of the spectrin alpha I/74-kilodalton tryptic  
 RL peptide."  
 RN Blood 78:1364-1372(1991).  
 [12]  
 RP VARIANT EL2 SER-45.  
 RX MEDLINE=89323468; PubMed=2568862;  
 RA Lecomte M.-C., Garbarz M., Grandchamp B., Feo C., Gautero H.,  
 RT Devaux I., Bourrier O., Galand C., D'Aurilio L., Galibert P.,  
 RA Sahr K.E., Forget B.G., Boivin P., Dhery D.;  
 RT "Sp alpha I/78: a mutation of the alpha I spectrin domain in a white  
 RL kindred with HE and HPP phenotypes."  
 RN Blood 74:1126-1133(1989).  
 [13]  
 RP VARIANT EL2/HPP PRO-207.  
 RX MEDLINE=92176375; PubMed=1541680;  
 RA Gallagher P.G., Tse W.T., Coetzer T., Lecomte M.-C., Garbarz M.,  
 RA Zarkowsky H.S., Baruchel A., Ballas S.K., Dhery D., Palek J.,  
 RT Forget B.G.;  
 RL "A common type of the spectrin alpha I 46-50a-kD peptide abnormality  
 RT in hereditary elliptocytosis and pyropoikilocytosis is associated  
 RT with a mutation distant from the proteolytic cleavage site. Evidence  
 RL for the functional importance of the triple helical model of  
 RT spectrin."  
 RN J. Clin. Invest. 89:892-898(1992).  
 [14]  
 RP VARIANT VAL-1857.  
 RX MEDLINE=93253053; PubMed=8486776;  
 RA Wilmette R., Marechal J., Morle L., Baklouti F., Philippe N.,  
 RT Kastally R., Kotula L., Delaunay J., Alloisio N.;  
 RL "Low expression allele alpha DELY of red cell spectrin is associated  
 RN with mutations in exon 40 (alpha V/41 polymorphism) and intron 45 and  
 with partial skipping of exon 46."  
 RN J. Clin. Invest. 91:2091-2096(1993).  
 [15]  
 RP VARIANT EL2 BARCELONA PRO-469.  
 RX MEDLINE=93372367; PubMed=8364215;  
 RA dalla Venezia N., Alloisio N., Forissier A., Denoroy L., Aymerich M.,  
 RT Vives-Corrons J.L., Besalduch J., Besson I., Delaunay J.;  
 RL "Elliptocytosis associated with the alpha 469 His-->Pro  
 mutation in spectrin Barcelona (alpha I/50-46b)."  
 RN Blood 82:1661-1665(1993).  
 [16]  
 RP VARIANT CAGLIARI GLY-2024.  
 RX MEDLINE=94043025; PubMed=8226774;  
 RA Sahr K.E., Coetzer T.L., Moy L.S., Derick L.H., Chishti A.H.,  
 RT Jarolim P., Lorenzo F., Miraglia del Giudice E., Iolascon A.,  
 RL Gallanello R.;  
 RN "Spectrin Cagliari: an Ala-->Gly substitution in helix 1 of beta  
 RT spectrin repeat 17 that severely disrupts the structure and self-  
 RL association of the erythrocyte spectrin heterodimer."  
 RN J. Biol. Chem. 268:22656-22662(1993).  
 [17]  
 RP VARIANT EL2 CULOZ VAL-46, AND VARIANT EL2 LYON PHE-49.  
 RX MEDLINE=90347052; PubMed=2384601;  
 RA Morle L., Roux A.-F., Alloisio N., Pothier B., Starck J., Denoroy J.,  
 RT Morle F., Rudigoz R.-C., Forget B.G., Delaunay J., Godet J.;  
 RL "Two elliptocytogenic alpha I/74 variants of the spectrin alpha I  
 RT domain. Spectrin Culoz (GGT-->GTT; alpha I 40 Gly-->Val) and  
 RN Spectrin Lyon (CTT-->TTT; alpha I 43 Leu-->Phe)."  
 RN J. Clin. Invest. 86:548-554(1990).  
 [18]  
 RP VARIANT EL2 JENDOUBA GLU-791.  
 RX MEDLINE=92345619; PubMed=1638030;  
 RA Alloisio N., Wilmette R., Morle L., Baklouti F., Marechal J.,  
 RT Ducluzeau M.-T., Denoroy L., Feo C., Forget B.G., Kastally R.,  
 RN Delaunay J.;  
 RL "Spectrin Jendouba: an alpha II/31 spectrin variant that is  
 RT associated with elliptocytosis and carries a mutation distant from  
 RL the dimer self-association site."  
 RN Blood 80:809-815(1992).  
 [19]  
 RP VARIANT EL2 TUNIS TRP-41.  
 RX MEDLINE=89323436; PubMed=2568861;  
 RA Morle L., Morle F., Roux A.-F., Godet J., Forget B.G., Denoroy L.,  
 RT Garbarz M., Dhery D., Kastally R., Delaunay J.;  
 RL "Spectrin Tunis (Sp alpha I/78), an elliptocytogenic variant, is due  
 RT to the CGG-->TGG codon change (Arg-->Trp) at position 35 of the  
 RL alpha I domain."  
 RN Blood 74:828-832(1989).  
 [20]  
 RP VARIANT EL2 GENOVA TRP-34.  
 RX MEDLINE=94250920; PubMed=8193371;  
 RA Perrotta S., del Giudice E.M., Alloisio N., Sciaratta G., Pinto L.,  
 RT Delaunay J., Cuttulo S., Iolascon A.;  
 RL "Wild elliptocytosis associated with the alpha 34 Arg-->Trp mutation  
 in spectrin Genova (alpha I/74)."  
 RN Blood 83:3346-3349(1994).  
 [21]  
 RP VARIANT EL2 ANASTASIA THR-45.  
 RX MEDLINE=95290423; PubMed=7772539;  
 RA Perrotta S., Iolascon A., de Angelis F., Pagano L., Colonna G.,  
 RT Cuttulo S., del Giudice E.M.;  
 RL "Spectrin Anastasia (alpha I/78): a new spectrin variant (alpha 45  
 Arg-->Thr) with moderate elliptocytogenic potential."  
 RN Br. J. Haematol. 89:933-936(1995).  
 CC -!- FUNCTION: Spectrin is the major constituent of the cytoskeletal  
 CC network underlying the erythrocyte plasma membrane. It associates  
 CC with band 4.1 and actin to form the cytoskeletal superstructure of  
 CC the erythrocyte plasma membrane.



## RESULT 48

PLE1\_HUMAN  
 ID\_PLE1\_HUMAN STANDARD; PRT; 4684 AA.  
 AC Q15149; Q16640;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HDI).  
 GN PLEC1  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX MEDLINE=96210632; PubMed=8633085;  
 RA Liu C.-G., Maercker C., Castanon M.J., Hauptmann R., Wiche G.;  
 RT "Human plectin: organization of the gene, sequence analysis, and  
 RT chromosome localization (8q24).";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283(1996).  
 [2]  
 RN SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND DISEASE.  
 RX MEDLINE=96312447; PubMed=8698233;  
 RA McLean W.H.I., Pulkkinen L., Smith F.J.D., Rugg E.L., Lane E.B.,  
 RA Bullrich F., Burgeson R.E., Amaro S., Hudson D.B., Owaribe K.,  
 RA McGrath J.A., McMillan J.R., Bady R.A.J., Leigh I.M., Christiano A.M.,  
 RA Utto J.;  
 RT "Loss of plectin causes epidermolysis bullosa with muscular dystrophy:  
 RT cDNA cloning and genomic organization.";  
 RL Genes Dev. 10:1724-1735(1996).  
 [3]  
 RN VARIANT MD-EBS 1003-GLN--ALA-1005 DEL.  
 RX MEDLINE=97049959; PubMed=8894687;  
 RA Pulkkinen L., Smith F.J.D., Shimizu H., Murata S., Yacita H.,  
 RA Hachisuka H., Nishikawa T., McLean W.H.I., Utto J.;  
 RT "Homozygous deletion mutations in the plectin gene (PLEC1) in patients  
 RT with epidermolysis bullosa simplex associated with late-onset  
 RT muscular dystrophy.";  
 RL Hum. Mol. Genet. 5:1539-1546(1996).  
 [4]  
 RN VARIANT MD-EBS LEU-429 INS.  
 RX MEDLINE=21090821; PubMed=11159198;  
 RA Bauer J.W., Rouan F., Kofler B., Reznicek G.A., Kornacker I.,  
 RA Muss W., Hametner R., Klausegger A., Huber A., Pohla-Gubo G.,  
 RA Wiche G., Utto J., Hinthner H.;  
 RT "A compound heterozygous one amino-acid insertion/nonsense mutation in  
 RT the plectin gene causes epidermolysis bullosa simplex with plectin  
 RT deficiency.";  
 RL Am. J. Pathol. 158:617-625(2001).  
 [5]  
 RN VARIANT EBS1 TRP-2110.  
 RX MEDLINE=21841370; PubMed=11851880;  
 RA Koss-Harnes D., Hoeyheim B., Anton-Lamprecht I., Gjesti A.,  
 RA Joergensen R.S., Jahnson F.L., Olaisen B., Wiche G.,  
 RA Gedde-Dahl T. Jr.;  
 RT "A site-specific plectin mutation causes dominant epidermolysis  
 RT bullosa simplex Ogna: two identical de novo mutations.";  
 RL J. Invest. Dermatol. 118:87-93(2002).  
 CC -1- FUNCTION: Interlinks intermediate filaments with microtubules and  
 CC microfilaments and anchors intermediate filaments to desmosomes or  
 CC hemidesmosomes. Could also bind muscle proteins such as actin to  
 CC membrane complexes in muscle. May be involved not only in the  
 CC crosslinking and stabilization of cytoskeletal intermediate  
 CC filaments network, but also in the regulation of their dynamics.  
 CC -1- SUBUNIT: Homodimer or homotetramer.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name=1;  
 CC IsoId=Q15149-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q15149-2; Sequence=VSP\_005030;  
 CC Name=3;

CC IsoId=Q15149-3; Sequence=VSP\_005030, VSP\_005031;  
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in  
 CC muscle, heart, placenta and spinal cord.  
 CC -1- DOMAIN: The N-terminus interacts with actin, the C-terminus with  
 CC vimentin, desmin, GFAP, cytokeratins, lamin B; whereas both the N-  
 CC and the C-terminus can bind integrin beta-4.  
 CC -1- PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM  
 CC INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY).  
 CC -1- DISEASE: Defects in PLEC1 are the cause of epidermolysis bullosa  
 CC simplex with muscular dystrophy (MD-EBS) [MIM:226670]; an  
 CC autosomal recessive disorder characterized by epidermal blister  
 CC formation at the level of the hemidesmosome and associated with  
 CC late-onset muscular dystrophy.  
 CC -1- DISEASE: Defects in PLEC1 are the cause of epidermolysis bullosa  
 CC simplex 1 (EBS1) [MIM:131950]; also called epidermolysis bullosa  
 CC simplex Ogna type. EBS1 is an autosomal dominant form of  
 CC epidermolysis bullosa simplex differentiated from the more  
 CC generalized form of Koebner [MIM:131900] and the localized form of  
 CC Weber and Cockayne [MIM:131800] by the occurrence of skin  
 CC bruising.  
 CC -1- SIMILARITY: Contains 1 actin-binding domain.  
 CC -1- SIMILARITY: Contains 2 calponin-homology (CH) domains.  
 CC -1- SIMILARITY: Contains 33 plectin repeats.  
 CC -1- SIMILARITY: Contains 4 spectrin repeats.  
 CC -1- SIMILARITY: Belongs to the plakin or cytolinker family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; Z54367; CAA91196.1; -;  
 CC EMBL; U53204; AAB05427.1; -;  
 CC EMBL; U53610; AAB05428.1; -;  
 CC EMBL; U53609; AAB05428.1; JOINED.  
 CC EMBL; X97053; CAA65765.1; -;  
 CC FIR; C59404; A59404.  
 CC HSP; Q01082; 1BKR.  
 CC Genew; HGNC:9069; PLEC1.  
 CC GK; Q15149; -;  
 CC MIM; 601282; -;  
 CC MIM; 226670; -;  
 CC MIM; 131950; -;  
 CC GO; GO:0008307; F:structural constituent of muscle; TAS.  
 CC InterPro; IPR001589; Actbind actnin.  
 CC InterPro; IPR001715; Calponin-like.  
 CC InterPro; IPR001101; Plectin repeat.  
 CC InterPro; IPR005326; S10\_plectin\_N.  
 CC InterPro; IPR002017; Spectrin.  
 CC Pfam; PF00307; CH; 2.  
 CC Pfam; PF00681; Plectin; 19.  
 CC Pfam; PF03501; S10\_plectin; 1.  
 CC ProDom; PD006662; S10\_plectin\_N; 1.  
 CC SMART; SM00033; CH; 2; S10\_plectin\_N; 1.  
 CC SMART; SM00250; PLEC; 34.  
 CC PROSITE; PS00019; ACTININ\_1; FALSE\_NEG.  
 CC PROSITE; PS00020; ACTININ\_2; FALSE\_NEG.  
 CC PROSITE; PS0021; CH; 2.  
 CC Coiled coil; Repeat; Structural protein; Cytoskeleton; Actin-binding;  
 CC Phosphorylation; Alternative splicing; Epidermolysis bullosa;  
 CC Disease mutation.  
 FT DOMAIN 1 1470 GLOBULAR 1.  
 FT DOMAIN 1471 2755 CENTRAL FIBROUS ROD DOMAIN.  
 FT DOMAIN 2756 4684 GLOBULAR 2.  
 FT DOMAIN 175 400 ACTIN-BINDING.  
 FT DOMAIN 179 282 CH 1.  
 FT DOMAIN 295 397 CH 2.  
 FT REPEAT 645 710 SPECTRIN 1.  
 FT REPEAT 740 824 SPECTRIN 2.  
 FT REPEAT 837 930 SPECTRIN 3.

```

REPEAT 1315 1415 SPECTRIN 4.
FT DOMAIN 1469 2756 COILED COIL (POTENTIAL).
FT REPEAT 2826 2863 PLECTIN 1.
FT REPEAT 2864 2901 PLECTIN 2.
FT REPEAT 2902 2939 PLECTIN 3.
FT REPEAT 2940 2977 PLECTIN 4.
FT REPEAT 2981 3015 PLECTIN 5.
FT REPEAT 3116 3153 PLECTIN 6.
FT REPEAT 3154 3191 PLECTIN 7.
FT REPEAT 3192 3229 PLECTIN 8.
FT REPEAT 3230 3267 PLECTIN 9.
FT REPEAT 3268 3305 PLECTIN 10.
FT REPEAT 3306 3343 PLECTIN 11.
FT REPEAT 3485 3522 PLECTIN 12.
FT REPEAT 3523 3560 PLECTIN 13.
FT REPEAT 3561 3598 PLECTIN 14.
FT REPEAT 3599 3636 PLECTIN 15.
FT REPEAT 3640 3674 PLECTIN 16.
FT REPEAT 3820 3857 PLECTIN 17.
FT REPEAT 3858 3895 PLECTIN 18.
FT REPEAT 3896 3933 PLECTIN 19.
FT REPEAT 3934 3971 PLECTIN 20.
FT REPEAT 3975 4008 PLECTIN 21.
FT REPEAT 4063 4100 PLECTIN 22.
FT REPEAT 4101 4138 PLECTIN 23.
FT REPEAT 4139 4176 PLECTIN 24.
FT REPEAT 4177 4214 PLECTIN 25.
FT REPEAT 4218 4252 PLECTIN 26.
FT REPEAT 4265 4305 PLECTIN 27.
FT REPEAT 4319 4355 PLECTIN 28.
FT REPEAT 4408 4445 PLECTIN 29.
FT REPEAT 4446 4483 PLECTIN 30.
FT REPEAT 4484 4521 PLECTIN 31.
FT REPEAT 4522 4559 PLECTIN 32.
FT REPEAT 4560 4597 PLECTIN 33.
FT DOMAIN 4625 4640 BINDING TO INTERMEDIATE FILAMENTS
FT MOD_RES 4539 (BY SIMILARITY).
FT REPEAT 1 174 4 X 4 AA TANDM REPEATS OF G-S-R-X.
FT VARSPLIC 1 174 PHOSPHORYLATION (BY CDC2)
MVAQLMPDQRLAIVELFREGVNVAKKRRPSLRHPHP
GVTNLQVMAASLRAGLVRETFAWCHFFWVLTNEGIAHL
ROYLHLPPEIVASLQVRPVPVAMPARTPHVCAVQGPL
GSPKRGPLPTEQLYRKLEEVSPETVPVPAITQKTLA
RQPEPAPAT -> WSGDAEVAIVSEVDSVNGSGSPSGD
TLFWNLGKTQRRRGSGGAGNSGVLDPAERAVIRIA (in
isoform 2 and isoform 3).
/FTid=VSP 005030.
Missing (in isoform 3).
/FTid=VSP 005031.
L > LL (in MD-EBS).
/FTid=VAR_011336.
Missing (in MD-EBS).
/FTid=VAR_011337.
R -> W (in EBS1).
/FTid=VAR_015817.

Query Match 40.0%; Score 38; DB 1; Length 4684;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 AKRGRKLVYDSAR 16
||| | | | | |
Db 1950 AKRQRLAEADAAR 1963

RESULT 49
VCOM_ADEL12
ID VCOM_ADEL12 STANDARD; PRT; 347 AA.
AC P36717;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)

Minor core protein (Protein V).
Human adenovirus type 12.
Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
NCBI_TaxID=28282;
[1]
SEQUENCE FROM N.A.
MEDLINE=94076430; PubMed=8254750;
Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;
"Nucleotide sequence of human adenovirus type 12 DNA: comparative
functional analysis.";
J. Virol. 68:379-389(1994).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X73487; CAAS1888.1; -.
PIR; S33939; S33939.
InterPro; IPR005608; Adeno_PV.
Pfam; PF03910; Adeno_PV; 1.
Core protein; Late protein.
SQ SEQUENCE 347 AA; 39369 MW; 3C338F62AA53027 CRC64;

Query Match 39.5%; Score 37.5; DB 1; Length 347;
Best Local Similarity 44.4%; Pred. NO. 73;
Matches 8; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 1 RIAGRGKLVYDSARHH 18
| : | | | | : |
Db 327 RVVQGRGRLI-LPSVRYH 343

RESULT 50
RS15_HAEIN
ID RS15_HAEIN STANDARD; PRT; 88 AA.
AC P44389;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-2003 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S15.
(RPSO-A OR RPS15-A OR H11328) AND (RPSO-B OR RPS15-B OR H11468).
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
NCBI_TaxID=727;
[1]
SEQUENCE FROM N.A.
STRAIN=RD / KW20 / ATCC 51907;
MEDLINE=95350630; PubMed=7542800;
Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
McKenney J.D., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.M.,
Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
Venter J.C.;
"Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
Science 269:496-512(1995).
-!- FUNCTION: This protein is one of the 16S ribosomal RNA binding
proteins (By similarity).
-!- SIMILARITY: Belongs to the S15P family of ribosomal proteins.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC EMBL; U32812; AAC22973.1; -.  
DR EMBL; U32825; AAC23117.1; -.  
DR PIR; H64116; H64116.  
DR HSSP; P05766; IA32.  
DR TIGR; H11328; -.  
DR TIGR; H11468; -.  
DR InterPro; IPR000589; Ribosomal\_S15.  
DR InterPro; IPR005290; Ribosomal\_S15\_b.  
DR Pfam; PF00312; Ribosomal\_S15; 1.  
DR ProDom; PD157043; RS15\_bact; 1.  
DR TIGRFAMs; TIGR00952; S15\_bact; 1.  
DR PROSITE; PS00362; RIBOSOMAL\_S15; 1.  
KW Ribosomal protein; rRNA-binding; Complete proteome.  
FT INIT MET 0 0 BY SIMILARITY.  
SQ SEQUENCE 88 AA; 10064 MW; 563BAD2B8B8A7043 CRC64;

Query Match 38.9%; Score 37; DB 1; Length 88;  
Best Local Similarity 59.3%; Pred. No. 20;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RIAKRGKLVY 12  
| : | | | | |  
Db 57 RMVSRKKLLDY 68

Search completed: March 4, 2004, 17:44:54  
Job time : 17.1935 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:41:21 ; Search time 18.5806 Seconds  
(without alignments)  
50.013 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_138\_155  
Perfect score: 95  
Sequence: 1 RIANKRKLVDYDSARHH 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pdp:\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pdp:\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pdp:\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pdp:\*  
5: /cgn2\_6/prodata/2/iaa/PCITUS COMB.pdp:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	404	4	US-08-630-915A-24
2	95	100.0	434	4	US-08-630-915A-22
3	95	100.0	451	1	US-08-435-454-4
4	95	100.0	451	1	US-08-652-972A-4
5	95	100.0	451	2	US-08-919-145-6
6	95	100.0	451	3	US-08-870-126-4
7	95	100.0	451	3	US-09-344-889-6
8	95	100.0	451	4	US-09-445-247-4
9	95	100.0	451	5	PCR-US96-06231A-4
10	45	47.4	1213	3	US-09-413-814-79
11	42	44.2	721	4	US-09-107-532A-6889
12	42	44.2	1219	4	US-09-252-991A-28840
13	40.5	42.6	298	4	US-09-107-532A-3726
14	39	41.1	67	4	US-09-543-681A-4534
15	39	41.1	95	4	US-09-328-352-8138
16	39	41.1	360	4	US-09-847-057-2
17	39	41.1	1282	4	US-09-543-681A-5419
18	38.5	40.5	724	4	US-09-562-737-25
19	38	40.0	258	4	US-09-252-991A-22623
20	38	40.0	344	4	US-09-134-001C-3158
21	38	40.0	421	4	US-09-252-991A-30366
22	38	40.0	476	4	US-09-489-039A-13646
23	38	40.0	489	4	US-09-252-991A-19960
24	38	40.0	566	1	US-08-419-078-2
25	38	40.0	566	1	US-08-726-883-2
26	38	40.0	615	2	US-08-272-253-7
27	38	40.0	615	5	PCR-US95-08565-7

28	38	40.0	887	4	US-09-252-991A-16679	Sequence 16679, A
29	38	40.0	990	4	US-09-252-991A-22792	Sequence 22792, A
30	38	40.0	1433	2	US-08-365-486A-21	Sequence 21, Appl
31	38	40.0	1433	3	US-09-123-708-4	Sequence 4, Appl
32	38	40.0	1433	3	US-09-123-624-4	Sequence 4, Appl
33	38	40.0	1433	3	US-08-880-342-21	Sequence 21, Appl
34	38	40.0	1434	2	US-08-365-486A-19	Sequence 19, Appl
35	38	40.0	1434	3	US-08-880-342-19	Sequence 19, Appl
36	38	40.0	1434	4	US-09-661-258-1	Sequence 1, Appl
37	38	40.0	1554	2	US-08-705-625-3	Sequence 3, Appl
38	38	40.0	1554	3	US-09-010-938-6	Sequence 6, Appl
39	38	40.0	1554	3	US-09-220-574-3	Sequence 3, Appl
40	37	38.9	88	4	US-09-732-210-1412	Sequence 1412, Ap
41	37	38.9	105	4	US-09-107-532A-5344	Sequence 5344, Ap
42	37	38.9	145	4	US-09-732-210-577	Sequence 577, App
43	37	38.9	163	4	US-09-134-001C-3257	Sequence 3257, Ap
44	37	38.9	361	4	US-09-134-000C-5619	Sequence 5619, Ap
45	37	38.9	365	2	US-08-827-190-5	Sequence 5, Appl
46	37	38.9	365	4	US-09-170-187-5	Sequence 5, Appl
47	37	38.9	381	4	US-09-252-991A-19173	Sequence 19173, A
48	37	38.9	387	4	US-09-180-109A-2	Sequence 2, Appl
49	37	38.9	387	4	US-09-180-109A-6	Sequence 6, Appl
50	37	38.9	387	4	US-09-180-109A-8	Sequence 8, Appl
51	37	38.9	404	4	US-08-252-991A-19166	Sequence 19166, A
52	37	38.9	415	4	US-08-252-991A-32170	Sequence 32170, A
53	37	38.9	424	4	US-09-252-991A-31825	Sequence 31825, A
54	37	38.9	459	4	US-09-134-001C-5466	Sequence 5466, Ap
55	37	38.9	466	4	US-09-252-991A-32286	Sequence 32286, A
56	37	38.9	535	4	US-09-252-991A-31973	Sequence 31973, A
57	37	38.9	662	3	US-08-779-814-5	Sequence 5, Appl
58	37	38.9	1277	4	US-09-397-885-3	Sequence 3, Appl
59	37	38.9	1277	4	US-09-969-362-3	Sequence 3, Appl
60	36.5	38.4	259	4	US-08-913-159-6	Sequence 6, Appl
61	36.5	38.4	269	2	US-08-424-641B-3	Sequence 3, Appl
62	36.5	38.4	269	2	US-08-820-980-3	Sequence 3, Appl
63	36.5	38.4	269	2	US-08-826-439-3	Sequence 3, Appl
64	36.5	38.4	269	4	US-08-913-159-4	Sequence 4, Appl
65	36.5	38.4	638	2	US-08-557-122A-38	Sequence 38, Appl
66	36.5	38.4	638	4	US-09-252-991A-11808	Sequence 38, Appl
67	36.5	38.4	1011	4	US-09-489-039A-8797	Sequence 11808, A
68	36	37.9	80	4	US-09-328-352-4702	Sequence 4702, Ap
69	36	37.9	82	4	US-09-107-532A-4820	Sequence 4820, Ap
70	36	37.9	88	4	US-09-732-210-1411	Sequence 1411, Ap
71	36	37.9	103	4	US-08-489-039A-13603	Sequence 13603, A
72	36	37.9	253	3	US-09-216-295-9	Sequence 9, Appl
73	36	37.9	253	3	US-09-216-295-10	Sequence 10, Appl
74	36	37.9	254	4	US-09-632-570-9	Sequence 9, Appl
75	36	37.9	254	4	US-09-632-570-10	Sequence 10, Appl
76	36	37.9	254	4	US-09-632-575-39	Sequence 39, Appl
77	36	37.9	254	4	US-09-632-575-40	Sequence 40, Appl
78	36	37.9	269	4	US-08-634-238-298	Sequence 298, App
79	36	37.9	274	4	US-09-252-991A-20389	Sequence 20389, A
80	36	37.9	281	4	US-09-423-439-44	Sequence 44, Appl
81	36	37.9	288	4	US-09-423-439-38	Sequence 38, Appl
82	36	37.9	290	4	US-09-252-991A-18490	Sequence 18490, A
83	36	37.9	305	4	US-09-252-991A-24732	Sequence 24732, A
84	36	37.9	321	4	US-09-107-532A-6016	Sequence 6016, Ap
85	36	37.9	326	4	US-09-543-681A-7915	Sequence 7915, Ap
86	36	37.9	339	4	US-09-252-991A-27733	Sequence 27733, A
87	36	37.9	406	4	US-09-252-991A-20630	Sequence 20630, A
88	36	37.9	408	4	US-09-328-352-5575	Sequence 5575, Ap
89	36	37.9	557	4	US-09-252-991A-22465	Sequence 22465, A
90	36	37.9	567	4	US-09-328-352-7224	Sequence 7224, Ap
91	36	37.9	583	4	US-09-328-352-5822	Sequence 5822, Ap
92	36	37.9	583	6	525658-4	Patent No. 525658
93	36	37.9	597	2	US-08-883-534-6	Sequence 6, Appl
94	36	37.9	597	3	US-09-204-764-6	Sequence 6, Appl
95	36	37.9	605	4	US-09-252-991A-18837	Sequence 18837, A
96	36	37.9	836	4	US-09-491-356C-21	Sequence 21, Appl
97	36	37.9	975	4	US-09-543-681A-5755	Sequence 5755, Ap
98	36	37.9	989	4	US-08-540-236-2137	Sequence 2137, Ap
99	36	37.9	989	4	US-09-786-240-6	Sequence 6, Appl
100	35.5	37.4	619	4		



101 35.5 37.4 724 4 US-09-562-737-23 Sequence 23, Appl  
102 35.5 37.4 916 4 US-09-252-991A-23637 Sequence 23637, A  
103 35.5 37.4 1441 4 US-09-252-991A-28143 Sequence 28143, A  
104 35 36.8 12 3 US-09-060-039-22 Sequence 22, Appl  
105 35 36.8 54 2 US-08-592-646A-15 Sequence 15, Appl  
106 35 36.8 54 4 US-09-165-422-15 Sequence 15, Appl  
107 35 36.8 83 4 US-09-134-000C-4372 Sequence 4372, Ap  
108 35 36.8 83 4 US-09-134-000C-4372 Sequence 4372, Ap  
109 35 36.8 96 4 US-09-621-976-4999 Sequence 4999, Ap  
110 35 36.8 139 4 US-09-107-532A-4424 Sequence 4424, Ap  
111 35 36.8 133 3 US-09-154-083-12 Sequence 12, Appl  
112 35 36.8 133 3 US-09-154-083-20 Sequence 20, Appl  
113 35 36.8 181 4 US-09-134-000C-6757 Sequence 6757, Ap  
114 35 36.8 218 4 US-09-252-991A-26105 Sequence 26105, A  
115 35 36.8 222 4 US-09-543-681A-6244 Sequence 6244, Ap  
116 35 36.8 267 4 US-09-134-001C-5042 Sequence 5042, Ap  
117 35 36.8 283 4 US-09-107-532A-6360 Sequence 6360, Ap  
118 35 36.8 318 4 US-09-107-532A-6731 Sequence 6731, Ap  
119 35 36.8 329 4 US-09-540-236-2444 Sequence 2444, Ap  
120 35 36.8 329 4 US-09-107-532A-7038 Sequence 7038, Ap  
121 35 36.8 329 4 US-09-543-681A-4790 Sequence 4790, Ap  
122 35 36.8 361 1 US-10-083-889-2 Sequence 1, Appl  
123 35 36.8 398 4 US-10-083-889-4 Sequence 4, Appl  
124 35 36.8 408 4 US-10-083-889-22 Sequence 22, Appl  
125 35 36.8 413 4 US-09-134-000C-3595 Sequence 3595, Ap  
126 35 36.8 419 4 US-09-252-991A-30457 Sequence 30457, A  
127 35 36.8 429 4 US-09-134-001C-4960 Sequence 4960, Ap  
128 35 36.8 458 4 US-09-252-991A-22614 Sequence 22614, A  
129 35 36.8 463 4 US-09-252-991A-24757 Sequence 24757, A  
130 35 36.8 464 4 US-09-297-468-2 Sequence 2, Appl  
131 35 36.8 485 4 US-09-479-645A-2 Sequence 2, Appl  
132 35 36.8 485 4 US-09-479-645A-4 Sequence 4, Appl  
133 35 36.8 608 4 US-09-489-039A-12785 Sequence 12785, A  
134 35 36.8 611 4 US-09-107-532A-4988 Sequence 4988, A  
135 35 36.8 640 1 US-09-177-165A-30 Sequence 30, Appl  
136 35 36.8 675 1 US-08-386-495-10 Sequence 10, Appl  
137 35 36.8 675 5 PCT-US96-02331-10 Sequence 10, Appl  
138 35 36.8 699 4 US-09-808-701A-34 Sequence 34, Appl  
139 35 36.8 715 4 US-09-808-701A-33 Sequence 33, Appl  
140 35 36.8 742 4 US-09-500-123-12 Sequence 12, Appl  
141 35 36.8 771 4 US-09-462-284-2 Sequence 2, Appl  
142 35 36.8 771 4 US-09-079-592-2 Sequence 2, Appl  
143 35 36.8 811 4 US-09-500-123-9 Sequence 9, Appl  
144 35 36.8 811 4 US-09-540-236-2412 Sequence 2412, Ap  
145 35 36.8 817 4 US-09-252-991A-20757 Sequence 20757, A  
146 35 36.8 871 4 US-09-500-123-7 Sequence 7, Appl  
147 35 36.8 895 4 US-09-489-039A-13127 Sequence 13127, A  
148 35 36.8 915 4 US-09-252-991A-24992 Sequence 24992, A  
149 35 36.8 944 4 US-09-134-000C-5578 Sequence 5578, Ap  
150 35 36.8 962 4 US-09-694-777A-24 Sequence 24, Appl

## ALIGNMENTS

RESULT 1  
US-08-630-915A-24  
; Sequence/24, Application US/08630915A  
; Patent No. 6309820  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: HOFFMAN, No. 6309820h  
; APPLICANT: KAY, Brian K.  
; APPLICANT: FOLKES, Dana M.  
; APPLICANT: MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; TITLE OF INVENTION: USING SAME  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630.915A  
FILING DATE: 03-APR-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 404 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-630-915A-24

Query Match 100.0%; Score 95; DB 4; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.1e-08; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RIAGRGKLVYDSARHH 18  
Db 104 RIAGRGKLVYDSARHH 121  
US-08-630-915A-24

RESULT 2  
US-08-630-915A-24  
; Sequence/22, Application US/08630915A  
; Patent No. 6309820  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: HOFFMAN, No. 6309820h  
; APPLICANT: KAY, Brian K.  
; APPLICANT: FOLKES, Dana M.  
; APPLICANT: MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; TITLE OF INVENTION: USING SAME  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630.915A  
FILING DATE: 03-APR-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-915A-22

Query Match 100.0%; Score 95; DB 4; Length 434;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
Db 136 RIAGRGKLVYDSARHH 156

RESULT 3  
US-08-435-454-4  
; Sequence 4, Application US/08435454  
; Patent No. 5605830  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; APPLICANT: Sakamuro, Daitoku  
; TITLE OF INVENTION: Murine and Human C-Myc Interacting  
; TITLE OF INVENTION: Protein and Uses Therefor  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,454  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: WST60USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-435-454-4

Query Match 100.0%; Score 95; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
Db 136 RIAGRGKLVYDSARHH 153

RESULT 4  
US-08-652-972A-4

; Sequence 4, Application US/0852972A  
; Patent No. 5723581  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; APPLICANT: Sakamuro, Daitoku  
; TITLE OF INVENTION: Murine and Human Box-Dependent  
; TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/652,972A  
; FILING DATE: 24-MAY-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,454  
; FILING DATE: 05-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: WST60USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-652-972A-4

Query Match 100.0%; Score 95; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAGRGKLVYDSARHH 18  
Db 136 RIAGRGKLVYDSARHH 153

RESULT 5  
US-08-919-145-6  
; Sequence 6, Application US/08919145  
; Patent No. 5958753  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; TITLE OF INVENTION: Bau, A Bin1 Interacting Protein, and  
; TITLE OF INVENTION: Uses Therefor  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/919,145
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,482
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST73AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-919-145-6
Query Match 100.0%; Score 95; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRGRKLVYDSARHH 18
Db 136 RIAGRGRKLVYDSARHH 153

RESULT 6
US-08-870-126-4
; Sequence 4, Application US/08870126
; Patent No. 6048702
; GENERAL INFORMATION:
; APPLICANT: Prendergast, George C.
; APPLICANT: Sakamuro, Daitoku
; TITLE OF INVENTION: Murine and Human Box-Dependent
; TITLE OF INVENTION: MYC-Interacting Protein (Bin1) and Uses Therefor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,126
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,454
; FILING DATE: 05-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,972
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST60CUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-870-126-4
Query Match 100.0%; Score 95; DB 3; Length 451;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRGRKLVYDSARHH 18
Db 136 RIAGRGRKLVYDSARHH 153

RESULT 7
US-09-344-889-6
; Sequence 6, Application US/09344889
; Patent No. 6140465
; GENERAL INFORMATION:
; APPLICANT: Prendergast, George C.
; TITLE OF INVENTION: Bau, A Bin1 Interacting Protein, and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,889
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,145
; FILING DATE:
; APPLICATION NUMBER: US 60/025,482
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST73AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-344-889-6
Query Match 100.0%; Score 95; DB 3; Length 451;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGRGRKLVYDSARHH 18
Db 136 RIAGRGRKLVYDSARHH 153

RESULT 8
US-09-445-247-4
; Sequence 4, Application US/09445247
; Patent No. 6410238
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy & Biology
```



; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6889:

SEQUENCE CHARACTERISTICS:

LENGTH: 721 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...721

SEQUENCE DESCRIPTION: SEQ ID NO: 6889:

US-09-107-532A-6889

Query Match 44.2%; Score 42; DB 4; Length 721;

Best Local Similarity 46.2%; Pred. No. 52;

Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 5 RGRKLVYDSARH 17

Db 90 RGKALDFQARH 102

RESULT 12

US-09-252-991A-28840

; Sequence 28840, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28840

; LENGTH: 1219

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28840

Query Match 44.2%; Score 42; DB 4; Length 1219;

Best Local Similarity 63.6%; Pred. No. 94;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 8 KLVYDSARH 18

Db 107 KLGDFQARH 117

RESULT 13

US-09-107-532A-3726

; Sequence 3726, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD/ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 3726:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 298 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...298

; SEQUENCE DESCRIPTION: SEQ ID NO: 3726:

US-09-107-532A-3726

Query Match 42.6%; Score 40.5; DB 4; Length 298;

Best Local Similarity 62.5%; Pred. No. 36;

Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RIA-KGRKLVYDSA 15

Db 180 RLAEKHGAKLVYDNA 195

RESULT 14

```

RESULT 16
US-09-847-057-2
; Sequence 2, Application US/09847057
; Patent No. 6509191
; GENERAL INFORMATION:
; APPLICANT: AGRINOMICS, LLC.
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A PAGODA PHENOTYPE IN PLANT
; FILE REFERENCE: PAGODA
; CURRENT APPLICATION NUMBER: US/09/847,057
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-847-057-2

```

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22623
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22623

Query Match      40.0%; Score 38; DB 4; Length 258;
Best Local Similarity 45.0%; Pred. No. 82;
Matches 9; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY      1 RIARGRKLV--DYDSARHH 18
        |||||
Db      56 RRQRRGRRLPLGDRSAGRHH 75

RESULT 20
US-09-134-001C-3158
; Sequence 3158, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3158
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3158

Query Match      40.0%; Score 38; DB 4; Length 344;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      6 GRKLVYDSA 15
        |||||
Db      10 GKXMDYKSA 19

RESULT 21
US-09-252-991A-30366
; Sequence 30366, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30366
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30366

Query Match      40.0%; Score 38; DB 4; Length 421;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      2 IAKRGKRLVDYDSARH 17
        |||||
Db      207 VAAAGRLHLDAGQRH 222

RESULT 22
US-09-489-039A-13646
; Sequence 13646, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13646
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13646

Query Match      40.0%; Score 38; DB 4; Length 476;
Best Local Similarity 44.4%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      1 RIARGRKRLVDYDSARHH 18
        |||||
Db      216 RYLRRFSVLDSEFSLRHH 233

RESULT 23
US-09-252-991A-19960
; Sequence 19960, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19960
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19960

Query Match      40.0%; Score 38; DB 4; Length 489;
Best Local Similarity 33.3%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      1 RIARGRKRLVDYDSARHH 18
        |||||

```

Db 195 QVVGEAGALDVDEAVHH 212

RESULT 24

US-08-419-078-2  
; Sequence 2, Application US/08419078  
; Patent No. 5587306  
; GENERAL INFORMATION:  
; APPLICANT: HAWKINS, PHILLIP R.  
; APPLICANT: SEILHAMER, JEFFREY J.  
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3330 HILLVIEW AVENUE  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/419,078  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUTHER, BARBARA J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PF0030 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0552  
; TELEFAX: 415-855-0572  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 566 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; IMMEDIATE SOURCE:  
; LIBRARY: NO. 5587306  
; CLONE: 9118  
US-08-419-078-2

Query Match 40.0%; Score 38; DB 1; Length 566;

Best Local Similarity 43.8%; Pred. No. 2e+02;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 25

US-08-726-883-2  
; Sequence 2, Application US/08726883  
; Patent No. 5676946  
; GENERAL INFORMATION:  
; APPLICANT: HAWKINS, PHILLIP R.  
; APPLICANT: SEILHAMER, JEFFREY J.  
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3330 HILLVIEW AVENUE  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,255  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Leary Ph.D., Kathryn  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: UPN-1795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 26

US-08-272-255-7  
; Sequence 7, Application US/08272255  
; Patent No. 5824859  
; GENERAL INFORMATION:  
; APPLICANT: Cashmore, Anthony R.  
; APPLICANT: Ahmad, Margaret  
; APPLICANT: Lin, Chentao  
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
; TITLE OF INVENTION: Using the Same  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,255  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Leary Ph.D., Kathryn  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: UPN-1795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 27

US-08-272-255-7  
; Sequence 7, Application US/08272255  
; Patent No. 5824859  
; GENERAL INFORMATION:  
; APPLICANT: Cashmore, Anthony R.  
; APPLICANT: Ahmad, Margaret  
; APPLICANT: Lin, Chentao  
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
; TITLE OF INVENTION: Using the Same  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,255  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Leary Ph.D., Kathryn  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: UPN-1795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/726,883  
; FILING DATE: 04-OCT-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/419,078  
; FILING DATE: 10-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUTHER, BARBARA J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PF0030 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-855-0572  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 566 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; IMMEDIATE SOURCE:  
; LIBRARY: NO. 5676946  
; CLONE: 9118  
US-08-726-883-2

Query Match 40.0%; Score 38; DB 1; Length 566;

Best Local Similarity 43.8%; Pred. No. 2e+02;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 26

US-08-272-255-7  
; Sequence 7, Application US/08272255  
; Patent No. 5824859  
; GENERAL INFORMATION:  
; APPLICANT: Cashmore, Anthony R.  
; APPLICANT: Ahmad, Margaret  
; APPLICANT: Lin, Chentao  
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
; TITLE OF INVENTION: Using the Same  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,255  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Leary Ph.D., Kathryn  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: UPN-1795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 27

US-08-272-255-7  
; Sequence 7, Application US/08272255  
; Patent No. 5824859  
; GENERAL INFORMATION:  
; APPLICANT: Cashmore, Anthony R.  
; APPLICANT: Ahmad, Margaret  
; APPLICANT: Lin, Chentao  
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
; TITLE OF INVENTION: Using the Same  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,255  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Leary Ph.D., Kathryn  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: UPN-1795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 28

US-08-272-255-7  
; Sequence 7, Application US/08272255  
; Patent No. 5824859  
; GENERAL INFORMATION:  
; APPLICANT: Cashmore, Anthony R.  
; APPLICANT: Ahmad, Margaret  
; APPLICANT: Lin, Chentao  
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
; TITLE OF INVENTION: Using the Same  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,255  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Leary Ph.D., Kathryn  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: UPN-1795  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

Qy 1 RIAGRGKLVYDSAR 16

Db 241 KIFQGESFVDYDGR 256

RESULT 29

US-08-272-255-7  
; Sequence 7, Application US/08272255  
; Patent No. 5824859  
; GENERAL INFORMATION:  
; APPLICANT: Cashmore, Anthony R.  
; APPLICANT: Ahmad, Margaret  
; APPLICANT: Lin, Chentao  
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
; TITLE OF INVENTION: Using the Same  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103



INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 615 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-272-255-7

Query Match 40.0%; Score 38; DB 2; Length 615;  
 Best Local Similarity 43.8%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 AKRGKLVYDSARHH 18  
 Db 312 APEGKRLRDEKARYH 327

## RESULT 27

PCT-US95-08565-7

Sequence 7, Application PC/TUS9508565  
 GENERAL INFORMATION:  
 APPLICANT: Cashmore, Anthony R.  
 APPLICANT: Ahmad, Margaret  
 APPLICANT: Lin, Chentao  
 TITLE OF INVENTION: Blue Light Photoreceptors and Methods of  
 TITLE OF INVENTION: Using the Same  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris  
 STREET: One Liberty Place, 46th floor  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/08565  
 FILING DATE:

CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/272,255  
 FILING DATE: 08-JUL-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Leary Ph.D., Kathryn  
 REGISTRATION NUMBER: 36,317  
 REFERENCE/DOCKET NUMBER: UPN-1795  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 568-3100  
 TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 615 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US95-08565-7

Query Match 40.0%; Score 38; DB 5; Length 615;  
 Best Local Similarity 43.8%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 AKRGKLVYDSARHH 18  
 Db 312 APEGKRLRDEKARYH 327

## RESULT 28

US-09-252-991A-16679  
 Sequence 16679, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 16679  
 LENGTH: 887  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-16679

Query Match 40.0%; Score 38; DB 4; Length 887;  
 Best Local Similarity 47.1%; Pred. No. 3.2e+02;  
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 IAKRGKLVYDSARHH 18  
 Db 253 LADRGASGVDDPRQH 269

## RESULT 29

US-09-252-991A-22792  
 Sequence 22792, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 22792  
 LENGTH: 990  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-22792

Query Match 40.0%; Score 38; DB 4; Length 990;  
 Best Local Similarity 46.2%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 5 RGKRLVYDSARH 17  
 Db 205 RGKRLVNDTKLH 217

## RESULT 30

US-08-365-486A-21  
 Sequence 21, Application US/08365486A  
 Patent No. 5834306  
 GENERAL INFORMATION:

APPLICANT: Webster, Keith A.  
 APPLICANT: Bishopric, Nanette H.  
 TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
 TITLE OF INVENTION: Therapeutic Constructs  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dehlinger & Associates

```
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/365,486A
; APPLICATION NUMBER: 08/365,486A
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-365-486A-21

Query Match 40.0%; Score 38; DB 2; Length 1433;
Best Local Similarity 62.5%; Pred. No. 5.5e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 2 IAKGRKLVLD--YDSA 15
Db 65 LAVNGRPLVLSYDSA 80

RESULT 31
US-09-123-708-4
; Sequence 4, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRAEDER, Juergen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2003
; CURRENT FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/553,503
; EARLIER FILING DATE: 1996-03-01
; EARLIER FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Cytomegalovirus
US-09-123-708-4

Query Match 40.0%; Score 38; DB 3; Length 1433;
Best Local Similarity 62.5%; Pred. No. 5.5e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 2 IAKGRKLVLD--YDSA 15
Db 65 LAVNGRPLVLSYDSA 80

RESULT 32
US-09-123-624-4
; Sequence 4, Application US/09123624
; Patent No. 6149936
; GENERAL INFORMATION:
; APPLICANT: SCHRAEDER, Juergen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2004
; CURRENT APPLICATION NUMBER: US/09/123,624
; CURRENT FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 08/553,503
; PRIOR FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: 4411402.8
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-123-624-4

Query Match 40.0%; Score 38; DB 3; Length 1433;
Best Local Similarity 62.5%; Pred. No. 5.5e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 2 IAKGRKLVLD--YDSA 15
Db 65 LAVNGRPLVLSYDSA 80

RESULT 33
US-08-880-342-21
; Sequence 21, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; APPLICANT: Murphy, Brian
; APPLICANT: Laderoute, Keith R.
; APPLICANT: Green, Christopher J.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/880,342
; APPLICATION NUMBER: 08/880,342
; FILING DATE: 23-JUN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB95/00996
; FILING DATE: 13-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/365,486
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
```

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1433 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-880-342-21

Query Match 40.0%; Score 38; DB 3; Length 1433;  
Best Local Similarity 62.5%; Pred. No. 5.5e+02;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKGRKLVLD--YDSA 15  
: |||||  
Db 65 LAVNGRPLVLDLSYDSA 80

RESULT 34  
US-08-365-486A-19  
Sequence 19, Application US/08365486A  
Patent No. 5834306  
GENERAL INFORMATION:

APPLICANT: Webster, Keith A.  
APPLICANT: Bishopric, Nanette H.  
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/365,486A  
FILING DATE: 23-DEC-1994  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8255-0018  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1434 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-365-486A-19

Query Match 40.0%; Score 38; DB 2; Length 1434;  
Best Local Similarity 62.5%; Pred. No. 5.5e+02;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKGRKLVLD--YDSA 15  
: |||||  
Db 65 LAVNGRPLVLDLSYDSA 80

RESULT 35  
US-08-880-342-19  
Sequence 19, Application US/08880342  
Patent No. 6216179  
GENERAL INFORMATION:  
APPLICANT: Webster, Keith A.

APPLICANT: Bishopric, Nanette H.  
APPLICANT: Murphy, Brian  
APPLICANT: Laderoute, Keith R.  
APPLICANT: Green, Christopher J.  
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/880,342  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IB95/00996  
FILING DATE: 13-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/365,486  
FILING DATE: 23-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8255-0018.30  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1434 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-880-342-19

Query Match 40.0%; Score 38; DB 3; Length 1434;  
Best Local Similarity 62.5%; Pred. No. 5.5e+02;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKGRKLVLD--YDSA 15  
: |||||  
Db 65 LAVNGRPLVLDLSYDSA 80

RESULT 36  
US-09-661-258-1  
Sequence 1, Application US/09661258  
Patent No. 6620616  
GENERAL INFORMATION:  
APPLICANT: Stuehr, Dennis J.  
APPLICANT: Adak, Subrata  
TITLE OF INVENTION: Nucleic Acids Encoding Nitric Oxide Synthase Variants  
FILE REFERENCE: 26473/04028  
CURRENT APPLICATION NUMBER: US/09/661,258  
CURRENT FILING DATE: 2000-09-13  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 1  
LENGTH: 1434  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-661-258-1

Query Match 40.0%; Score 38; DB 4; Length 1434;

Best Local Similarity 62.5%; Pred. No. 5.5e+02;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKRGKLVLD--YDSA 15  
Db 65 LAVNGRPLVDLSYDSA 80

## RESULT 37

US-08-705-625-3  
; Sequence 3, Application US/08705625  
; Patent No. 5908756  
; GENERAL INFORMATION:  
; APPLICANT: Snyder, Solomon H.  
; APPLICANT: Jaffrey, Samie R.  
; TITLE OF INVENTION: Protein Inhibitor of Neuronal Nitric  
; TITLE OF INVENTION: Oxide Synthase  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff  
; STREET: 1001 G Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001-4597  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/705,625  
; FILING DATE: 30-AUG-1996  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.57071  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1554 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORGANISM: Rattus norvegicus  
US-08-705-625-3

Query Match 40.0%; Score 38; DB 2; Length 1554;  
Best Local Similarity 62.5%; Pred. No. 6e+02;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKRGKLVLD--YDSA 15  
Db 65 LAVNGRPLVDLSYDSA 80

## RESULT 38

US-09-010-998-6  
; Sequence 6, Application US/09010998  
; Patent No. 6103872  
; GENERAL INFORMATION:  
; APPLICANT: Snyder, Solomon  
; APPLICANT: Jaffrey, Samie  
; APPLICANT: Snowman, Adele  
; APPLICANT: Eliasson, Mikael  
; TITLE OF INVENTION: CAPON, a protein that binds  
; TITLE OF INVENTION: neuronal nitric oxide synthase  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/010,998  
; FILING DATE: 22-JAN-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A  
; REGISTRATION NUMBER: 32141  
; REFERENCE/DOCKET NUMBER: 01107.73424  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1554 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6103872e  
US-09-010-998-6

Query Match 40.0%; Score 38; DB 3; Length 1554;  
Best Local Similarity 62.5%; Pred. No. 6e+02;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKRGKLVLD--YDSA 15  
Db 65 LAVNGRPLVDLSYDSA 80

## RESULT 39

US-09-220-574-3  
; Sequence 3, Application US/09220574  
; Patent No. 6168926  
; GENERAL INFORMATION:  
; APPLICANT: Snyder, Solomon H.  
; APPLICANT: Jaffrey, Samie R.  
; TITLE OF INVENTION: Protein Inhibitor of Neuronal Nitric  
; TITLE OF INVENTION: Oxide Synthase  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff  
; STREET: 1001 G Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001-4597  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/220,574  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/705,625  
; FILING DATE: 30-AUG-1996  
; ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.57071  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Rattus norvegicus  
US-09-220-574-3

Query Match 40.0%; Score 38; DB 3; Length 1554;  
Best Local Similarity 62.5%; Pred. No. 6e+02; 3; Indels 2; Gaps 1;  
Matches 10; Conservative 1; Mismatches 3

Qy 2 IAKRGRKLVLD--YDSA 15  
Db 65 LAVNGRPLVLSYDSA 80

## RESULT 40

US-09-732-210-1412  
Sequence 1412, Application US/09732210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Mittanck, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yennie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21(15036)B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 1753  
SEQ ID NO 1412  
LENGTH: 88  
TYPE: PRT  
ORGANISM: Haemophilus influenzae  
US-09-732-210-1412

Query Match 38.9%; Score 37; DB 4; Length 88;  
Best Local Similarity 58.3%; Pred. No. 37;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RIAKRGRKLVYD 12  
Db 57 RWSRRRKLDY 68

## RESULT 41

US-09-732-210-5344  
Sequence 5344, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A. Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts

COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Atiniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5344:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...105  
SEQUENCE DESCRIPTION: SEQ ID NO: 5344:  
US-09-107-532A-5344

Query Match 38.9%; Score 37; DB 4; Length 105;  
Best Local Similarity 46.7%; Pred. No. 45;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 KRGRKLVYDSARH 18  
Db 29 KKNTKLCPHDSNRH 43

## RESULT 42

US-09-732-210-577  
Sequence 577, Application US/09732210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Mittanck, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yennie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21(15036)B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 1753  
SEQ ID NO 577  
LENGTH: 145  
TYPE: PRT  
ORGANISM: Staphylococcus carnosus  
US-09-732-210-577

Query Match 38.9%; Score 37; DB 4; Length 145;  
Best Local Similarity 50.0%; Pred. No. 64;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGRKLVYDSARH 17

Db 118 KQKKLFVYGGAEH 131

RESULT 43

US-09-134-001C-3257  
; Sequence 3257, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3257  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3257

Query Match 38.9%; Score 37; DB 4; Length 163;

Best Local Similarity 50.0%; Pred. No. 73;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGRKLVYDSARH 17

Db 136 KQKKLFVYGGAEH 149

RESULT 44

US-09-134-000C-5619  
; Sequence 5619, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5619  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5619

Query Match 38.9%; Score 37; DB 4; Length 361;

Best Local Similarity 63.6%; Pred. No. 1.8e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 KRGRKLVYDS 14

Db 10 KRGRTLKEYES 20

RESULT 45

US-08-827-190-5  
; Sequence 5, Application US/08827190  
; Patent No. 5853367  
; GENERAL INFORMATION:

APPLICANT: Rather, Philip N.  
TITLE OF INVENTION: Methods For Screening For Antimicrobials  
Utilizing aarC And Compositions Thereof  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/827,190  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: CASE-02443  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-827-190-5

Query Match 38.9%; Score 37; DB 2; Length 365;

Best Local Similarity 60.0%; Pred. No. 1.8e+02;

Matches 9; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 4 KRGRKLVYDSARH 18

Db 116 ERIRQV-DSARH 128

RESULT 46

US-09-170-187-5  
; Sequence 5, Application US/09170187  
; Patent No. 6383745  
; GENERAL INFORMATION:  
; APPLICANT: Rather, Philip N.  
; TITLE OF INVENTION: Methods For Screening For Antimicrobials  
; Utilizing aarC And Compositions Thereof  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/170,187  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/827,190  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:







QY 61 YKDLKNFLSAVKVWHESSKRVSETLOEYSSWDGHEELKAIWNNDLWEDYEKLADQ 120  
 Db 61 YKDLKNFLSAVKVWHESSKRVSETLOEYSSWDGHEELKAIWNNDLWEDYEKLADQ 120  
 QY 121 AVRTMEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 180  
 Db 121 AVRTMEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 180  
 QY 181 TVPFEDLNQELLELPILYNSRICGYVTIFONISNLRDVFYR 221  
 Db 181 TVPFEDLNQELLELPILYNSRICGYVTIFONISNLRDVFYR 221

## RESULT 2

Q86VV0 ID Q86VV0 PRELIMINARY; PRT; 565 AA.  
 AC Q86VV0;  
 DT 01-JUN-2003 (TREMELrel. 24, Created)  
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Breast cancer associated protein BRAP1.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RK MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Beak S.A., McEwan P.J., McKernan K.J., Walek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalka U., Smalls D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RA Strausberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC047686; AAH47686.1; -  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR SMART; SM00721; BAR; 1.  
 SQ SEQUENCE 565 AA; 61875 MW; 81D1FB99A7707EA3 CRC64;

Query Match 99.1%; Score 1114.5; DB 4; Length 565;  
 Best Local Similarity 99.1%; Pred. No. 2.7e-63;  
 Matches 221; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MAEGKAGAGLFAKQVOKFSAQEKVLQKLGKAVETKDERFEQASNFYQQQAEHGKL 60  
 Db 1 MAEGKAGAGLFAKQVOKFSAQEKVLQKLGKAVETKDERFEQASNFYQQQAEHGKL 60

QY 61 YKDLKNFLSAVKVWHESSKRVSETLOEYSSWDGHEELKAIWNNDLWEDYEKLADQ 120  
 Db 61 YKDLKNFLSAVKVWHESSKRVSETLOEYSSWDGHEELKAIWNNDLWEDYEKLADQ 120  
 QY 121 AVRTMEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 179  
 Db 121 AVRTMEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 180  
 QY 180 QTVFEDLNQELLELPILYNSRICGYVTIFONISNLRDVFYR 221  
 Db 181 QTVFEDLNQELLELPILYNSRICGYVTIFONISNLRDVFYR 222

## RESULT 3

Q9UBW5 ID Q9UBW5 PRELIMINARY; PRT; 565 AA.  
 AC Q9UBW5;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Breast cancer associated protein BRAP1.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RK MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Beak S.A., McEwan P.J., McKernan K.J., Walek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalka U., Smalls D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RA Strausberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC047686; AAH47686.1; -  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR SMART; SM00721; BAR; 1.  
 SQ SEQUENCE 565 AA; 61902 MW; 28D1FB89A7779860 CRC64;

Query Match 98.8%; Score 1111.5; DB 4; Length 565;  
 Best Local Similarity 99.1%; Pred. No. 4.2e-63;  
 Matches 220; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MAEGKAGAGLFAKQVOKFSAQEKVLQKLGKAVETKDERFEQASNFYQQQAEHGKL 60  
 Db 1 MAEGKAGAGLFAKQVOKFSAQEKVLQKLGKAVETKDERFEQASNFYQQQAEHGKL 60  
 QY 61 YKDLKNFLSAVKVWHESSKRVSETLOEYSSWDGHEELKAIWNNDLWEDYEKLADQ 120  
 Db 61 YKDLKNFLSAVKVWHESSKRVSETLOEYSSWDGHEELKAIWNNDLWEDYEKLADQ 120  
 QY 121 AVRTMEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 179  
 Db 121 AVRTMEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 180  
 QY 180 QTVFEDLNQELLELPILYNSRICGYVTIFONISNLRDVFYR 221

Db 181 QTVFEDLNQELLELPILYNSRIGCYVTIFQNISLNRDVFYR 222

## RESULT 4

ID Q80313 PRELIMINARY; PRT; 524 AA.  
AC Q80313;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to bridging integrator 2.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC044473; AAH44473.1; -.  
DR GO; GO:0006897; P: endocytosis; IEA.  
DR GO; GO:0007268; P: synaptic transmission; IEA.  
DR InterPro; IPR003005; Amphiphysin.  
DR InterPro; IPR006632; BAR.  
DR Pfam; PF03114; BAR; 1.  
DR PRINTS; PR01251; AMPHIPHYSIN.  
DR SMART; SM00721; BAR; 1.  
SQ SEQUENCE 524 AA; 59280 MW; F4A2988732D153FF CRC64;

Query Match 69.8%; Score 782.5; DB 13; Length 524;

Best Local Similarity 65.7%; Pred. No. 3.7e-42;

Matches 151; Conservative 34; Mismatches 36; Indels 9; Gaps 2;

QY 1 MARGK-----AGGAAGLFAKQVKFSAQEKVLOKLGKAVETKDERFEQASNFYQ 52  
DB 1 MADGKLTGNIWNIGAGILAKFKQSNRAQEKVLOKLGKAVETKDERFEQASNFYQ 60  
QY 53 QQAEGHKLKDLNFKNPLSVKVMHSSKRVSETLQRIYSSEWDGHEELKAIYVNNDDLWEDYEEKLAD 112  
DB 61 QQTDDGLRLYKDKVAYNAVKVMHSSKRVSETLQRIYSSEWDGHEELKAIYVNNDDLWEDYEEKLAD 120  
QY 113 YEEKLADQAVRTMEIYVAQFSIKERIAKGRKLVYDVSARHLEAVQNA-KDEAKTAKAEEFNK 171  
DB 121 YEEKLADQAVRTMEIYVAQFSIKERIAKGRKLVYDVSARHLEAVQNA-KDEAKTAKAEEFNK 180  
QY 172 AEEFNKAQTVFEDLNQELLELPILYNSRIGCYVTIFQNISLNRDVFYR 221  
DB 181 AEEFNKAQTVFEDLNQELLELPILYNSRIGCYVTIFQNISLNRDVFYR 230

## RESULT 5

ID Q9BTH3 PRELIMINARY; PRT; 439 AA.  
AC Q9BTH3;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Similar to bridging integrator 1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kainine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,

RA Phelan M., Farmer A.;  
RT Cloning of human full-length CDS in BD Creator(TM) System Donor  
vector.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

CC -P- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

DR EMBL; BC004101; AAH04101.1; -.

DR EMBL; BT006865; AAP35511.1; -.

DR GO; GO:0006897; P: endocytosis; IEA.

DR GO; GO:0007268; P: synaptic transmission; IEA.

DR InterPro; IPR003005; Amphiphysin.

DR InterPro; IPR006632; BAR.

DR InterPro; IPR004148; BAR\_dom.

DR InterPro; IPR000875; Cectropin.

DR InterPro; IPR001452; SH3.

DR Pfam; PF03114; BAR; 1.

DR Pfam; PF00018; SH3; 1.

DR PRINTS; PR01251; AMPHIPHYSIN.

DR PRINTS; PR00452; SH3DOMAIN.

DR PRODOM; PD000066; SH3; 1.

DR SMART; SM00721; BAR; 1.

DR SMART; SM00326; SH3; 1.

DR PROSITE; PS00268; CECTROPIN; 1.

DR PROSITE; PS50002; SH3; 1.

KW SH3 domain.

SQ SEQUENCE 439 AA; 48286 MW; 350E428C8E198136 CRC64;

Query Match 60.7%; Score 683; DB 4; Length 439;

Best Local Similarity 61.0%; Pred. No. 6.7e-36;

Matches 136; Conservative 32; Mismatches 53; Indels 2; Gaps 2;

QY 1 MAE-KGAGAGLFAKQVKFSAQEKVLOKLGKAVETKDERFEQASNFYQQAEGHK 59  
DB 1 MAEMSGKGTAGKIASNVQKLTAKQEKVLOKLGKAVETKDERFEQASNFYQQAEGHK 60  
QY 60 LYKDLNFKNPLSVKVMHSSKRVSETLQRIYSSEWDGHEELKAIYVNNDDLWEDYEEKLAD 119  
DB 61 LQKDLRTYLSVKVMHSSKRVSETLQRIYSSEWDGHEELKAIYVNNDDLWEDYEEKLAD 120  
QY 120 QAVRTMEIYVAQFSIKERIAKGRKLVYDVSARHLEAVQNA-KDEAKTAKAEEFNK 178  
DB 121 QALLTMDTVLGQFPDIKSIKAKGRKLVYDVSARHLEAVQNA-KDEAKTAKAEEFNK 180  
QY 179 AQTVEEDLNQELLELPILYNSRIGCYVTIFQNISLNRDVFYR 221  
DB 181 AOKVFEEMVNDVQLELPSLWNSRVGYVNTFQSTAGLEENFKH 223

## RESULT 6

ID Q8WH9 PRELIMINARY; PRT; 490 AA.  
AC Q8WH9;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Amphiphysin IIB-1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Kim K.-C., Kim T.-S., Kang K.-H., Choi K.-H.;  
RT "Amphiphysin IIB-1, a novel splicing variant of amphiphysin II,  
regulates p73b function through protein-protein interactions.";  
RL Oncogene 0:0-0(2002).  
DR EMBL; AF411606; AAL38509.1; -.  
DR GO; GO:0006897; P: endocytosis; IEA.  
DR GO; GO:0007268; P: synaptic transmission; IEA.  
DR InterPro; IPR003005; Amphiphysin.  
DR InterPro; IPR006632; BAR.  
DR InterPro; IPR004148; BAR\_dom.  
DR InterPro; IPR001452; SH3.

DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 SQ SEQUENCE 490 AA; 53167 MW; 8F50F36F7B6E9690 CRC64;

Query Match 60.7%; Score 683; DB 4; Length 490;  
 Best Local Similarity 61.0%; Pred. No. 7.6e-36;  
 Matches 136; Conservative 32; Mismatches 53; Indels 2; Gaps 2;

QY 1 MAE-GKAGAGLFAKQVQKFSRAQKVLQKLGKAVETKDERFEQASNFYQQQAEHGK 59  
 DB 1 MAEMSGKGVAGKIASNVQKGLTRAQKVLQKLGKADETKDEQFECVQNFKNQLTEGTR 60  
 QY 60 LYKDLKNFLSAVKVMHSSKSEVSTLQEIYSSWDGHEELKAIVWNDLLWEDYEKLAD 119  
 DB 61 LOKLRTYLSVKAMHSEKSKINCELOEVYEPDPMPGRDEANKIAENNDLLWMDYHQKQVD 120  
 QY 120 QAVRTMEIYVAFSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKABEEFNK 178  
 DB 121 QALLTMDTYLGQFPDIKSRIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKABEEFNK 180  
 QY 179 AQTFVEDLNQELFELPILYNSRIGCVYTFIFONISNLRDVYR 221  
 DB 181 AQKVFEMNVDLQBELPSLWNSRVGFYVNTFQSIAGLEENFHK 223

RESULT 7  
 Q7ZWP5 PRELIMINARY; PRT; 478 AA.

AC Q7ZWP5  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Similar to bridging integrator 1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 ON NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (FEJ-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046852; AAH46852.1; -.  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR InterPro; IPR004452; SH3.  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 SQ SEQUENCE 478 AA; 53445 MW; 042A1A94D59095A8 CRC64;

Query Match 58.9%; Score 662.5; DB 13; Length 478;  
 Best Local Similarity 58.6%; Pred. No. 1.5e-34;  
 Matches 130; Conservative 35; Mismatches 56; Indels 1; Gaps 1;

QY 1 MAEKGAGAGLFAKQVQKFSRAQKVLQKLGKAVETKDERFEQASNFYQQQAEHGK 60  
 DB 1 MAELGKGVAGKIASNVQKGLTRAQKVLQKLGKADETKDEQFECVQNFKNQLTEGSK 60  
 QY 61 YKDLKNFLSAVKVMHSSKSEVSTLQEIYSSWDGHEELKAIVWNDLLWEDYEKLADQ 120

DB 61 QKDLRAYLATVKAMHEASKKLTCELLDVEYDVPDMPGREETNKIAENNDLLWTDYHQKLVQD 120  
 QY 121 AVRTMEIYVAFSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKABEEFNK 179  
 DB 121 ALLTMDTYLGQFPDIKSRIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKABEEFNK 180  
 QY 180 QTFVEDLNQELFELPILYNSRIGCVYTFIFONISNLRDVYR 221  
 DB 181 QKVFEMNVDLQBELPSLWNSRVGFYVNTFQSIAGLEENFHK 222

RESULT 8  
 Q8BXH2 PRELIMINARY; PRT; 250 AA.

AC Q8BXH2  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Similar to MYC box dependent interacting protein 1.  
 GN AMPH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The PANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RL Nature 420:563-573 (2002)."  
 DR EMBL; AK047144; BAC32971.1; -.  
 DR MGD; MGI:103574; Amph.  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR SMART; SM00721; BAR; 1.  
 SQ SEQUENCE 250 AA; 29125 MW; 568A3D55CC6C37CE CRC64;

Query Match 55.6%; Score 625.5; DB 11; Length 250;  
 Best Local Similarity 55.2%; Pred. No. 1.6e-32;  
 Matches 117; Conservative 49; Mismatches 45; Indels 1; Gaps 1;

QY 11 GLFAKQVQKFSRAQKVLQKLGKAVETKDERFEQASNFYQQQAEHGKLYKLNFLSA 70  
 DB 7 GIFAKVQKELNRAQKVLQKLGKADETKDEQFECVQNFKNQLTEGRLGYLAA 66  
 QY 71 VKVMHSSKSEVSTLQEIYSSWDGHEELKAIVWNDLLWEDYEKLADQAVRTMEIYVA 130  
 DB 67 IKGQEASMKLTESLHEVYEPDPMPGRDEAVKMGKCDVLWEDFHQKLVGSLTLDYVLG 126  
 QY 131 QFSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKABEEFNKQTFVEDLNQ 189  
 DB 127 QFPDIKNRIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKABEEFNKQTFVEDLNQ 186  
 QY 190 LLEELPILYNSRIGCVYTFIFONISNLRDVYR 221  
 DB 187 LQBELPSLWNSRVGFYVNTFKNVSSLEAKFHK 218

RESULT 9  
 Q7TQF7 PRELIMINARY; PRT; 686 AA.

AC Q7TQF7  
 DT 01-OCT-2003 (Tremblrel. 25, Created)  
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalka U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.  
 RA EMBL; BC054718; AAF54718.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 686 AA; 75013 MW; D292E24653A442A5 CRC64;  
 Query Match 55.6%; Score 625.5; DB 11; Length 686;  
 Best Local Similarity 55.2%; Pred. No. 5.2e-32;  
 Matches 117; Conservative 49; Mismatches 45; Indels 1; Gaps 1;  
 QY 11 GLFAKQVQKFSRAQKVLQKLGKAVETKDRFEQASNFYQQQAEGHKLKLNFLSA 70  
 Db 7 GIFAKNVQKLNRAQKVLQKLGKADETKDSQFEYVQNFKQAEGRQLQRELGLAA 66  
 QY 71 VKVMHSSKRVSTLQEIYSSWDGHEELKAIWNNDLLWEDYEKLADQAVRTMEIYVA 130  
 Db 67 IKGMQASMKLTESLHEVYDPWYGRDVGMVGEKCDVLMEDFHQKLVGSLLLDTYLG 126  
 QY 131 QPSEIKERIAKGRKLVYDSARHLEAVQNAK-KDEAKTAKAEFEFNKAQTVFEDLNQE 189  
 Db 127 QFPDIKNRIAKSRKLVYDSARHLEALQSSKRDERSISKAEEEFQAKQVFEFNVD 186  
 QY 190 LLEELPILNSRIGCVYTFIFONISNLRDVYR 221  
 Db 187 LQELPSLWSRRVGFYNTFNKVSLEAKFKH 218  
 RESULT 10  
 ID Q8NFK3 PRELIMINARY; PRT; 385 AA.  
 AC Q8NFK3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amphiphysin I variant NT2 (fragment).  
 GN AMPH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Lung;  
 RA Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; BC034376; AAF34376.1; --  
 DR GO; GO:0007268; P:synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR003017; Amphiphysin\_1.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR dom.  
 DR InterPro; IPR001452; SH3-  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO1251; AMPHIPHYSIN.  
 DR PRINTS; PRO0452; SH3DOMAIN.  
 DR ProDom; PD003208; Amphiphysin\_1; 1.

RP SEQUENCE FROM N.A.  
 RX MEDLINE=22017878; PubMed=12023042;  
 RA Terada Y., Tsutsui K., Sano K., Hosoya O., Ohtsuki H., Tokunaga A.,  
 RA Teutsui K.;  
 RT "Novel splice variants of amphiphysin I are expressed in retina.";  
 RL FEBS Lett. 519:185-190(2002).  
 DR EMBL; AF498097; AAM44811.1; --  
 DR GO; GO:0006897; P:endocytosis; IEA.  
 DR GO; GO:0007268; P:synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR003017; Amphiphysin\_1.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PRO1251; AMPHIPHYSIN.  
 DR ProDom; PD003208; Amphiphysin\_1; 1.  
 DR SMART; SM00721; BAR; 1.  
 FT NON\_TER 385 385  
 SQ SEQUENCE 385 AA; 43753 MW; 4B9AB0A3136711AA CRC64;  
 Query Match 55.2%; Score 620.5; DB 4; Length 385;  
 Best Local Similarity 54.7%; Pred. No. 5.6e-32;  
 Matches 116; Conservative 49; Mismatches 46; Indels 1; Gaps 1;  
 QY 11 GLFAKQVQKFSRAQKVLQKLGKAVETKDRFEQASNFYQQQAEGHKLKLNFLSA 70  
 Db 7 GIFAKNVQKLNRAQKVLQKLGKADETKDSQFEYVQNFKQAEGRQLQRELGLAA 66  
 QY 71 VKVMHSSKRVSTLQEIYSSWDGHEELKAIWNNDLLWEDYEKLADQAVRTMEIYVA 130  
 Db 67 IKGMQASMKLTESLHEVYDPWYGRDVGMVGEKCDVLMEDFHQKLVGSLLLDTYLG 126  
 QY 131 QPSEIKERIAKGRKLVYDSARHLEAVQNAK-KDEAKTAKAEFEFNKAQTVFEDLNQE 189  
 Db 127 QFPDIKNRIAKSRKLVYDSARHLEALQSSKRDERSISKAEEEFQAKQVFEFNVD 186  
 QY 190 LLEELPILNSRIGCVYTFIFONISNLRDVYR 221  
 Db 187 LQELPSLWSRRVGFYNTFNKVSLEAKFKH 218  
 RESULT 11  
 ID Q8N4G0 PRELIMINARY; PRT; 695 AA.  
 AC Q8N4G0;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amphiphysin (Stiff-Man syndrome with breast cancer 128kd  
 DE autoantigen).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Lung;  
 RA Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; BC034376; AAF34376.1; --  
 DR GO; GO:0007268; P:synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR003017; Amphiphysin\_1.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR dom.  
 DR InterPro; IPR001452; SH3-  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO1251; AMPHIPHYSIN.  
 DR PRINTS; PRO0452; SH3DOMAIN.  
 DR ProDom; PD003208; Amphiphysin\_1; 1.







GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:29:55 ; Search time 16.8917 Seconds  
(without alignments)  
2762.588 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_1\_221

Perfect score: 1125

Sequence: 1 MARGKAGGAGLFAKQVQKK.....ICGYTIFQNLNLRDVFYR 221

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/ECT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	689	61.2	434	9	US-09-879-957-22
2	683	60.7	482	15	US-10-116-275-238
3	671.5	59.7	451	14	US-10-123-807-4
4	559.5	49.7	404	9	US-09-879-957-24
5	181.5	16.1	284	15	US-10-369-493-2410
6	179.5	16.0	265	15	US-10-369-493-1519
7	160.5	14.3	452	15	US-10-369-493-3119
8	158.5	14.1	489	15	US-10-369-493-2512
9	155.5	13.8	351	15	US-10-369-493-3387
10	154.5	13.7	482	15	US-10-369-493-1690
11	148.5	13.2	265	15	US-10-369-493-12681
12	144	12.8	237	15	US-10-369-493-2408
13	142.5	12.7	420	15	US-10-369-493-13382
14	139.5	12.4	474	14	US-10-032-585-7847
15	118.5	10.5	650	15	US-10-104-047-3636

Sequence 3308, Ap  
Sequence 1061, Ap  
Sequence 7646, Ap  
Sequence 5874, Ap  
Sequence 5875, Ap  
Sequence 1586, Ap  
Sequence 16, Appl  
Sequence 1095, Ap  
Sequence 164, App  
Sequence 2, Appl  
Sequence 162, App  
Sequence 103, App  
Sequence 4, Appl  
Sequence 1682, Ap  
Sequence 21875, A  
Sequence 6, Appl  
Sequence 4, Appl  
Sequence 6, Appl  
Sequence 14, Appl  
Sequence 21887, A  
Sequence 4573, Ap  
Sequence 449, App  
Sequence 22080, A  
Sequence 2621, Ap  
Sequence 10617, A  
Sequence 7611, Ap  
Sequence 166, App  
Sequence 19, Appl  
Sequence 108, App  
Sequence 1950, Ap

16 114.5 10.2 853 15 US-10-320-797-3308  
17 113 10.0 1005 15 US-10-369-493-1061  
18 113 10.0 1881 14 US-10-032-585-7646  
19 112.5 10.0 1138 15 US-10-369-493-5874  
20 112.5 10.0 1138 15 US-10-369-493-5875  
21 111.5 9.9 1790 15 US-10-369-493-1586  
22 111 9.9 600 15 US-10-367-687-16  
23 110.5 9.8 1169 15 US-10-369-493-1095  
24 110 9.8 1938 14 US-10-171-311-164  
25 110 9.8 1945 10 US-09-927-597-2  
26 110 9.8 1972 14 US-10-171-311-162  
27 110 9.8 1972 15 US-10-341-434-103  
28 110 9.8 1979 10 US-09-927-597-4  
29 109.5 9.7 944 15 US-10-369-493-1682  
30 109.5 9.7 1225 15 US-10-369-493-21875  
31 109.5 9.7 2139 9 US-09-727-384-6  
32 109.5 9.7 2139 14 US-10-023-219-4  
33 109 9.7 402 14 US-10-284-400-6  
34 109 9.7 403 14 US-10-284-400-14  
35 109 9.7 1170 15 US-10-369-493-21887  
36 108.5 9.6 824 15 US-10-108-260A-4573  
37 108.5 9.6 1043 15 US-10-310-154-449  
38 108.5 9.6 1679 15 US-10-369-493-22080  
39 107.5 9.6 936 15 US-10-104-047-2621  
40 107.5 9.6 1045 9 US-09-815-242-10617  
41 107.5 9.6 1948 14 US-10-032-585-7611  
42 107 9.5 860 15 US-10-080-334-166  
43 107 9.5 975 14 US-10-080-608A-19  
44 107 9.5 975 15 US-10-370-685-108  
45 106.5 9.5 1312 15 US-10-369-493-1950

#### ALIGNMENTS

#### RESULT 1

US-09-879-957-22  
; Sequence 22, Application US/09879957  
; Patent No. US20020034755A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; HOFFMAN, No. US20020034755A1h  
; KAY, Brian K.  
; FOWLKES, Dana M.  
; McCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; USING SAME  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/879,957  
; FILING DATE: 13-Jun-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/630,915  
; FILING DATE: 03-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-174  
; TELECOMMUNICATION INFORMATION:



TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 TELE: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 434 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
 US-09-879-957-22

Query Match 61.2%; Score 689; DB 9; Length 434;  
 Best Local Similarity 61.9%; Pred. No. 6.2e-48;  
 Matches 138; Conservative 30; Mismatches 53; Indels 2; Gaps 2;  
 QY 1 MAE-GKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQAEGHK 59  
 Db 1 MAEWSKGVTAGKIASNVQKLTQAEKVLQKLGKADETKDEQECVQNFNKLTEGTR 60  
 QY 60 LYKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEKLD 119  
 Db 61 LQDLRTYLSVKAMHEASKKLSCELVYEPWPGDEANKIAENNDLLWMDYHQKLV 120  
 QY 120 QAVTMEIYVAQSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKAEFEFNK 178  
 Db 121 QALLTMDTYLGQFPDIKSRIKGRKLVYDYSARHYESLQTAKKDEAKIAKEELIK 180  
 QY 179 AQTVEFDLNOELLELPILYNSRGICVVTIFQNISNLRDVFYR 221  
 Db 181 AQKVFEMVNDLQBELPSLWNSRVGFYNTFQSIAGLEENPHK 223

## RESULT 2

US-10-116-275-238  
 Sequence 238, Application US/10116275  
 Publication No. US20030211476A1  
 GENERAL INFORMATION:  
 APPLICANT: Elan Pharmaceutical Technology  
 APPLICANT: O'Mahony, Daniel J.  
 APPLICANT: Brayden, David  
 APPLICANT: Byrne, Daragh  
 APPLICANT: Lambkin, Imelda  
 APPLICANT: Higgins, Lisa  
 TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
 TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors  
 FILE REFERENCE: E1067/20087  
 CURRENT FILING DATE: 2002-10-04  
 NUMBER OF SEQ ID NOS: 349  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 238  
 LENGTH: 482  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-116-275-238

Query Match 60.7%; Score 683; DB 15; Length 482;  
 Best Local Similarity 61.0%; Pred. No. 2.2e-47;  
 Matches 136; Conservative 32; Mismatches 53; Indels 2; Gaps 2;  
 QY 1 MAE-GKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQAEGHK 59  
 Db 1 MAEWSKGVTAGKIASNVQKLTQAEKVLQKLGKADETKDEQECVQNFNKLTEGTR 60  
 QY 60 LYKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEKLD 119  
 Db 61 LQDLRTYLSVKAMHEASKKLSCELVYEPWPGDEANKIAENNDLLWMDYHQKLV 120  
 QY 120 QAVTMEIYVAQSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKAEFEFNK 178  
 Db 121 QALLTMDTYLGQFPDIKSRIKGRKLVYDYSARHYESLQTAKKDEAKIAKEELIK 180

QY 179 AQTVEFDLNOELLELPILYNSRGICVVTIFQNISNLRDVFYR 221  
 Db 181 AQKVFEMVNDLQBELPSLWNSRVGFYNTFQSIAGLEENPHK 223

## RESULT 3

US-10-123-807-4  
 Sequence 4, Application US/10123807  
 Publication No. US20030166021A1  
 GENERAL INFORMATION:  
 APPLICANT: Wistar Institute of Anatomy & Biology  
 Prendergast, George C.  
 Sakamuro, Daitoku  
 TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Howson and Howson  
 STREET: Spring House Corporate Cntr, P O Box 457  
 CITY: Spring House  
 STATE: Pennsylvania  
 COUNTRY: USA  
 ZIP: 19477  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/123,807  
 FILING DATE: 16-Apr-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/445,247  
 FILING DATE: 03-Dec-1999  
 APPLICATION NUMBER: US 08/870,126  
 FILING DATE: 08-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bak, Mary E.  
 REGISTRATION NUMBER: 31,215  
 REFERENCE/DOCKET NUMBER: WST60DPCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-540-9200  
 TELEFAX: 215-540-5618  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 451 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 US-10-123-807-4

Query Match 59.7%; Score 671.5; DB 14; Length 451;  
 Best Local Similarity 61.5%; Pred. No. 1.7e-46;  
 Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;  
 QY 10 AGLFAKQVKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQAEGHKLYKDLKNFLS 69  
 Db 8 AGKIASNVQKLTQAEKVLQKLGKADETKDEQECVQNFNKLTEGTRLQDLRTYLA 67  
 QY 70 AVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEKLDQAVTMEIYV 129  
 Db 68 SVKAMHEASKKLSCELVYEPWPGDEANKIAENNDLLWMDYHQKLVQALLTMDTYL 127  
 QY 130 AQSFIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKAEFEFNKAQTVFEDLNQ 188  
 Db 128 CQFPDIKSRIKGRKLVYDYSARHYESLQTAKKDEAKIAKEELIKAKQVFEMV 187  
 QY 189 ELLELPILYNSRGICVVTIFQNISNLRDVFYR 221  
 Db 188 DQBELPSLWNSRVGFYNTFQSIAGLEENPHK 220

## RESULT 4

US-03-879-957-24  
; Sequence 24, Application US/09879957  
; Patent No. US20020034755A1  
; GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.  
HOFFMAN, No. US20020034755A1h  
KAY, Brian K.  
FOWLKES, Dana M.  
McCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/879,957

FILING DATE: 13-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915

FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mirock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 404 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 24:

US-03-879-957-24

## Query Match

Best Local Similarity 49.7%; Score 559.5; DB 9; Length 404;

Matches 107; Conservative 30; Mismatches 43; Indels 1; Gaps 1;

QY 42 RFEGASNFYQQQAEGHKLKYLKFNFLSAVKVMHESKRYSETLQEIYSEWDGHEELKA 101

DB 8 RFEGCVNFNKLQTEGTRLOKDLRTYLASVKAMHEASKKLNCELOEVYEPDFGRDEANK 67

QY 102 IVMNNDLLWDEYEKLDAQVRTWEIYVAQFSIKERIAKRGKLYDVSARHLEAVQN 161

DB 58 IAEENDLLWMDYHOKGVQDQALLMTYVIGQFDIKRIAKRGKLYDVSARHLYESLQT 127

QY 162 A-KKDEAKTAKAEFEFNKAQTVFEDLNQELBELPLIYNSRICGYVTIFQNTSNLRDVFY 220

DB 128 AKKXDEAKIAKAEELIKAKVPEENVNVDLQELPLSNVRGVYNTFQSIAGLEENFH 187

QY 221 R 221

DB 188 K 188

## RESULT 5

US-10-369-493-2410

; Sequence 2410, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 2410

LENGTH: 284

TYPE: PRT

ORGANISM: Schizosaccharomyces pombe

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(284)

OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-2410

## Query Match

Best Local Similarity 16.1%; Score 181.5; DB 15; Length 284;

Matches 51; Conservative 33; Mismatches 87; Indels 13; Gaps 5;

QY 28 VLOKLGKAVETKDEREQASNFYQQQAEGHKLKYLKFNFLSAVKVMHESKRYSETLQOE 87

DB 37 VMKGTGHVTVDRFETETERRYRTWESAKKLQKQKGYLDALRAMTASQTRIANTIDA 96

QY 88 IYSEWDGHEELKAIYMNNDLLWDEYEE---KLDAQVRTWEI-----YVAQFSIKERI 139

DB 97 FYGDA--GSKDGVSAVYRQ--VVEDLDADTVKELDGPFTTVDLPISRFCSYFPDINAAI 152

QY 140 AKRGKLYDVSARHLE-AVQNAKQDEAKTAKAEFEFNKAQTVFEDLNQELBELPLIY 198

DB 153 TKRNHKLDDHDAWRKVKLVDPKSPNDTTLKLPRTKEAAMAKEVYETLNNQLVSELPOLI 212

QY 199 NSRI 202

DB 213 ALRV 216

## RESULT 6

US-10-369-493-1519

; Sequence 1519, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 1519

LENGTH: 285

TYPE: PRT

ORGANISM: Saccharomyces cerevisiae

US-10-369-493-1519





OTHER INFORMATION: unsure at all xaa locations  
US-10-369-493-2408

Query Match 12.8%; Score 144; DB 15; Length 237;  
Best Local Similarity 26.4%; Pred. No. 0.00052;  
Matches 46; Conservative 27; Mismatches 75; Indels 26; Gaps 5;  
QY 40 DERFEQASNFYQQAEGHKLKLDKNFLSAVVMHSSKRVSETLQEIY--SSEWDG-- 95  
DB 1 DREFTERRVRYTWESAAKLOKEAGYLDALRAMTASQTRXANTIDAFYGDAGSKDGV 60  
QY 96 -----HEELKAIWNNDLLWEDYEKLDAQVRT-----MEIYVAQFSEIKERIAKGR 144  
DB 61 AYRQVVEDLDA-----DTVKELDXLXRTTVXGPIRSCYFPDINAATKRNH 110  
QY 145 KLVDYDSARHLE-AVQNAKDEAKTAKAEFEFNKAQTVFEDLNQELLELPIL 197  
DB 111 KLLDHDAWRAKVLVDKPSNDITKLPRTKEAAMAKVEYETLNKXTGSELPOL 164

## RESULT 13

US-10-369-493-13382  
Sequence 13382, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

PRIOR FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 13382

LENGTH: 420

TYPE: PRT

ORGANISM: Aspergillus nidulans

NAME/KEY: unsure

LOCATION: (1)-(420)

OTHER INFORMATION: unsure at all xaa locations  
US-10-369-493-13382

Query Match 12.7%; Score 142.5; DB 15; Length 420;  
Best Local Similarity 23.1%; Pred. No. 0.0014;  
Matches 58; Conservative 40; Mismatches 104; Indels 49; Gaps 5;

QY 15 KQVQKFSRAQEKVLQKLGKAVETKDRFEQASNFYQQAEGHKLKLDKNFLSAV--- 71  
DB 4 XGFQKSVIRAPQTPFKARFNIGDNTKDPVYTDARRRFELEKETKCLHDESKKYFDVNGR 63  
QY 72 -----KVMHSSKRVSETLQEIYSEWDGHEELKAIYWN 105  
DB 64 FRPDTFLHSGXPXAGMLDQHEFSKAMTELKPISGRASPDPSTVTIEGNEGTRACEE 123  
QY 106 NDLWEYEEKLADQ-----AVRTMEIYVAQFSEIKERIAKGRKLVYDSARH 154  
DB 124 YEVIQDLKEALAPLELMIDSRVSPQAQLLEVIKA-----IRKVAVRDHKKLDYDRHN 179  
QY 155 HLEAVONAK---KDEAKTAKAEFEFNKAQTVFEDLNQELLELPILYNSRIGCVTFIQ 210  
DB 180 TLKKLQDKQKSLKDEKALYKAGDVEQATQETNLYNDLLKDLKLPKLFALAEFTIRPLFQ 239  
QY 211 NISNL-RDVFY 220  
DB 240 SFYTMQLNVFY 250

## RESULT 14

US-10-032-585-7847  
Sequence 7847, Application US/10032585  
Publication No. US20030180953A1  
GENERAL INFORMATION:

APPLICANT: Terry, Roemer D.  
APPLICANT: Bo, Jiang

APPLICANT: Charles, Boone

APPLICANT: Howard, Bussey

TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

FILE REFERENCE: 10182-005-999

CURRENT APPLICATION NUMBER: US/10/032,585

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 8000

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7847

LENGTH: 474

TYPE: PRT

ORGANISM: Candida albicans

US-10-032-585-7847

Query Match 12.4%; Score 139.5; DB 14; Length 474;  
Best Local Similarity 21.5%; Pred. No. 0.0029;  
Matches 53; Conservative 46; Mismatches 103; Indels 45; Gaps 6;

QY 15 KQVQKFSRAQEKVLQKLGKAVETKDRFEQASNFYQQAEGHKLKLDKNFLSAVKNM 74  
DB 38 KGFKGVLRAPQTPMRQKFNWGEITQDAVYLDARRRFELEKETKCLSESKKYFNANVM 97  
QY 75 HESSKRVSETLQEIYS-----SEWDGHEELKAIYWN-NDLLWED--- 112  
DB 98 LDEQIDFAKAAEIIYKPISGRLSDPSATVPEDNPGTIEASESYQAVVVKDLDKTLKPLEL 157  
QY 113 YEEKLADQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNAK---KDEAK 168  
DB 158 IEKGIIVEPAQELKIIQA-----IRKMSVKRDHKQLDLDRHKNFSKYESKERTVKDBEK 213  
QY 169 TAKAEFEFNKAQTVFEDLNQELLELPIL-----YNSRIGCVYTFIQNIS 213  
DB 214 MFSQAQAEVIAQQEYDYNDLLKNEPLVLFQMSDFIKPLFVSYFYNQLNFIYLYTRME 273  
QY 214 NLRDVFY 220  
DB 274 ELKIPYF 280

## RESULT 15

US-10-104-047-3636  
Sequence 3636, Application US/10104047  
Publication No. US20030236392A1  
GENERAL INFORMATION:

APPLICANT: HELIX RESEARCH INSTITUTE

TITLE OF INVENTION: No. US20030236392A1el full length cDNA

FILE REFERENCE: H1-A0105

CURRENT APPLICATION NUMBER: US/10/104,047

CURRENT FILING DATE: 2002-03-25

PRIOR APPLICATION NUMBER:

PRIOR FILING DATE:

NUMBER OF SEQ ID NOS: 4096

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3636

LENGTH: 650

TYPE: PRT

ORGANISM: Homo sapiens

US-10-104-047-3636

Query Match 10.5%; Score 118.5; DB 15; Length 650;  
Best Local Similarity 21.2%; Pred. No. 0.022;  
Matches 40; Conservative 40; Mismatches 76; Indels 33; Gaps 7;

QY 18 QKFSRAQEKVLQKLGKAVETKDRFEQASNFYQQAEGHKLKLDKNFLSAVKNMHS 77  
DB 277 QBEELREQEKIRKQEKQKQWQBEELREQQKMGKEQE---EKMWRQSKRLRQEKELREQ 333



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:38:36 ; Search time 42.3871 Seconds  
(without alignments)  
133.987 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_138\_155

Perfect score: 95

Sequence: 1 RIAKGRKLVYDSARHH 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database :

SPREMBL\_25.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mmc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertibrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	439	4	Q9BTH3
2	95	100.0	490	4	Q8WH9
3	95	100.0	564	4	Q9UKN4
4	95	100.0	565	4	Q9UBW5
5	95	100.0	565	4	Q86VW0
6	91	95.8	478	13	Q7ZNP5
7	90	94.7	524	13	Q80313
8	89	93.7	250	11	Q8BXH2
9	89	93.7	385	4	Q8NFL3
10	89	93.7	686	11	Q7TQF7
11	89	93.7	695	4	Q8N4G0
12	89	92.6	461	5	Q21004
13	82	65.3	602	5	Q3T092
14	53	55.8	273	3	Q8X0D7
15	53	55.8	414	5	Q8WQ54
16	49.5	52.1	130	16	Q8X2J8

17	48	50.5	169	17	Q9HQS1	Q9hgs1 halobacteri
18	47	49.5	424	10	O64842	O64842 arabidopsis
19	47	49.5	425	10	Q39229	Q39229 arabidopsis
20	47	49.5	441	10	Q9L2A1	Q9leal arabidopsis
21	47	49.5	578	10	Q9ZST7	Q9zat7 arabidopsis
22	47	49.5	578	10	Q9ZST6	Q9zat6 arabidopsis
23	47	49.5	578	10	Q9LV77	Q9lv77 arabidopsis
24	47	49.5	578	10	Q9LFU1	Q9lfu1 arabidopsis
25	46	48.4	1163	16	Q89JES	Q89jes bradyrhizob
26	45.5	47.9	83	16	Q83JK1	Q83jk1 shigella fl
27	45.5	47.9	596	4	Q9H7T7	Q9h7t7 homo sapien
28	45.5	47.9	742	4	Q8NA04	Q8na04 homo sapien
29	45.5	47.9	763	4	Q96R86	Q96r86 homo sapien
30	45.5	47.9	813	4	Q8NE21	Q8ne21 homo sapien
31	45.5	47.9	836	4	Q8H2F5	Q8h2f5 mus sapien
32	45	47.4	128	9	Q852V9	Q852v9 mycobacteri
33	45	47.4	160	5	Q9UOX9	Q9uox9 leishmania
34	45	47.4	171	11	Q9D7W8	Q9d7w8 mus musculu
35	45	47.4	171	11	Q810A3	Q810a3 mus musculu
36	45	47.4	190	9	Q3ZXL0	Q3zx10 bacterioph
37	45	47.4	524	16	Q8KD94	Q8kd94 chlorobium
38	44	46.3	263	16	Q7WDT3	Q7wdt3 bordetella
39	44	46.3	263	16	Q7W2T5	Q7w2t5 bordetella
40	44	46.3	297	16	Q7V8X1	Q7v8x1 bordetella
41	44	46.3	524	2	Q8GDV2	Q8gdv2 heliobacill
42	43.5	45.8	282	16	Q8ZN24	Q8zn24 salmonella
43	43.5	45.8	282	16	Q8ZAL1	Q8zal1 salmonella
44	43	45.3	349	16	Q7WFG2	Q7wfg2 bordetella
45	43	45.3	349	16	Q7VTS5	Q7vts5 bordetella
46	43	45.3	413	11	Q8BLU6	Q8blu6 mus musculu
47	43	45.3	819	11	Q8BUT2	Q8but2 mus musculu
48	43	45.3	859	11	Q8BWW6	Q8bww6 mus musculu
49	43	45.3	863	11	Q8BZ98	Q8bz98 mus musculu
50	43	45.3	869	11	Q9OXL9	Q9oxl9 rattus norv
51	42	44.2	130	16	Q8E3D2	Q8e3d2 streptococ
52	42	44.2	130	16	Q8DXR3	Q8dxr3 streptococ
53	42	44.2	252	17	O28300	O28300 archaeoglob
54	42	44.2	257	17	Q979L7	Q979l7 thermoplas
55	42	44.2	368	10	Q7XIE8	Q7xie8 oryza sativ
56	42	44.2	463	16	O53770	O53770 mycobacteri
57	42	44.2	463	16	Q7UIR1	Q7uir1 mycobacteri
58	42	44.2	487	16	Q988V2	Q988v2 rhizobium l
59	42	44.2	578	16	Q8E4E2	Q8e4e2 coxiella bu
60	42	44.2	652	12	Q91AX4	Q91ax4 human papil
61	42	44.2	756	2	Q8KLH2	Q8klh2 rhizobium e
62	42	44.2	1193	16	Q9HZE5	Q9hze5 pseudomonas
63	42	44.2	1289	2	Q93IR7	Q93ir7 salmonella
64	42	44.2	1289	16	Q8ZRL2	Q8zrl2 salmonella
65	42	44.2	1957	5	Q04010	Q04010 onchocerca
66	41.5	43.7	181	10	Q940D5	Q940d5 cucumis mel
67	41.5	43.7	341	5	Q9UAM1	Q9uam1 leptomonas
68	41.5	43.7	361	5	O96425	O96425 leptomonas
69	41.5	43.7	593	16	Q826F7	Q826f7 streptomyc
70	41.5	43.7	731	10	Q39547	Q39547 cucumis mel
71	41.5	43.7	763	11	Q92299	Q92299 mus musculu
72	41.5	43.7	813	11	Q8C9X6	Q8c9x6 mus musculu
73	41	43.2	196	5	Q9XZT2	Q9xzt2 drosophila
74	41	43.2	196	5	Q9W588	Q9w588 drosophila
75	41	43.2	234	2	O52596	O52596 agrobacteri
76	41	43.2	237	16	O86353	O86353 mycobacteri
77	41	43.2	237	16	Q77231	Q77231 mycobacteri
78	41	43.2	338	5	O18341	O18341 bombyx mori
79	41	43.2	402	5	Q22968	Q22968 caenorhabdi
80	41	43.2	402	5	O58864	O58864 pyrococcus
81	41	43.2	457	16	Q7VHM6	Q7vhw6 heliobacte
82	41	43.2	469	16	Q8RAM6	Q8ram6 thermococ
83	41	43.2	469	16	Q8PVE1	Q8pve1 methanocarc
84	41	43.2	626	12	Q9DIH5	Q9dih5 human papil
85	41	43.2	629	10	Q9S7M1	Q9s7m1 oryza sativ
86	41	43.2	933	16	Q83GL1	Q83gl1 tropheryma
87	41	43.2	939	16	Q83HL6	Q83hl6 tropheryma
88	41	43.2	985	5	O01590	O01590 caenorhabdi
89	41	43.2	1001	2	Q9JP94	Q9jp94 rhodocyclu





DR InterPro; IPR004148; BAR\_dom.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00019; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR PRODOM; PD000066; SH3; 1.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 SQ SEQUENCE 490 AA; 53167 MW; 8F50F36F7B6E9690 CRC64;

Query Match 100.0%; Score 95; DB 4; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAKGRKLVYDSARHH 18  
 |||||  
 Db 139 RIAKGRKLVYDSARHH 156

## RESULT 3

Q9UKN4 PRELIMINARY; PRT; 564 AA.  
 AC Q9UKN4

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Bridging integrator-2.  
 GN BIN2.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

[1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20366138; PubMed=10903846;  
 RX Ge K., Prendergast G.C.;

RT "Bin2, a Functionally Nonredundant Member of the BAR Adaptor Gene  
 Family";  
 RL Genomics 67:210-220(2000).

DR EMBL; AF146531; AAD54227.1; -.  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.

DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR SMART; SM00721; BAR; 1.

SQ SEQUENCE 564 AA; 61747 MW; 3CB3791A56CE53CB CRC64;

Query Match 100.0%; Score 95; DB 4; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAKGRKLVYDSARHH 18  
 |||||  
 Db 138 RIAKGRKLVYDSARHH 155

## RESULT 4

Q9UBW5 PRELIMINARY; PRT; 565 AA.  
 AC Q9UBW5

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Breast cancer associated protein BRAP1.  
 GN BRAP1.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

[1]

## RP SEQUENCE FROM N.A.

RA Miki Y., Saito H.;  
 RT "Genomic structure and chromosome location of the BRAP1 gene";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB032710; BAA88125.1; -.  
 DR EMBL; AB032699; BAA88125.1; JOINED.  
 DR EMBL; AB032700; BAA88125.1; JOINED.  
 DR EMBL; AB032701; BAA88125.1; JOINED.  
 DR EMBL; AB032702; BAA88125.1; JOINED.

DR EMBL; AB032703; BAA88125.1; JOINED.  
 DR EMBL; AB032704; BAA88125.1; JOINED.  
 DR EMBL; AB032705; BAA88125.1; JOINED.  
 DR EMBL; AB032706; BAA88125.1; JOINED.  
 DR EMBL; AB032707; BAA88125.1; JOINED.

DR EMBL; AB032708; BAA88125.1; JOINED.  
 DR EMBL; AB032709; BAA88125.1; JOINED.  
 DR EMBL; AB032698; BAA88108.1; -.  
 DR Genew; HGNC:1053; BIN2.

DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.

DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.

DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR SMART; SM00721; BAR; 1.

SQ SEQUENCE 565 AA; 61902 MW; 28D1FB89A7779860 CRC64;

Query Match 100.0%; Score 95; DB 4; Length 565;

Best Local Similarity 100.0%; Pred. No. 2.7e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAKGRKLVYDSARHH 18

|||||  
 Db 138 RIAKGRKLVYDSARHH 155

## RESULT 5

Q86VV0

ID Q86VV0 PRELIMINARY; PRT; 565 AA.

AC Q86VV0;

DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Breast cancer associated protein BRAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

[1]

## RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RC MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loguillano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047686; AAH47686.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 565 AA; 61875 MW; B1D1FB99A7707EA3 CRC64;

Query Match 100.0%; Score 95; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 138 RIAKGRKLVYDSARHH 155

RESULT 6
Q7ZWPS PRELIMINARY; PRT; 478 AA.
AC Q7ZWPS;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DE Similar to bridging integrator 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S.; Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046852; AAH46852.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR02251; AMPHIPHYSIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
SQ SEQUENCE 478 AA; 53445 MW; 042A1A94D59095A8 CRC64;

Query Match 95.8%; Score 91; DB 13; Length 478;
Best Local Similarity 94.4%; Pred. No. 1.1e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 138 RIAKGRKLVYDSARHH 155

RESULT 7
Q80313 PRELIMINARY; PRT; 524 AA.
AC Q80313;
DT 01-JUN-2003 (Tremblrel. 24, Created)

```

```

DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to bridging integrator 2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044473; AAH44473.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 524 AA; 59280 MW; F4A2988732D153FF CRC64;

Query Match 94.7%; Score 90; DB 13; Length 524;
Best Local Similarity 88.9%; Pred. No. 1.7e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 146 RIAKGRKLVYDSARHH 163

RESULT 8
Q8BXH2 PRELIMINARY; PRT; 250 AA.
AC Q8BXH2;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to MYC box dependent interacting protein 1.
GN AMPH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT 'Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK047144; BAC32971.1; -.
DR MGD; MGI:103574; Amph.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 250 AA; 29125 MW; 568A3D55CC6C37CE CRC64;

Query Match 93.7%; Score 89; DB 11; Length 250;
Best Local Similarity 94.4%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 134 RIAKGRKLVYDSARHH 151

```

```

RESULT 9
Q8NFK3 PRELIMINARY; PRT; 385 AA.
ID AC Q8NFK3
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Amphiphsin I variant NT2 (fragment).
GN AMPH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22017878; PubMed=12023042;
RA Terada Y., Tsutsui K., Sano K., Hosoya O., Ohtsuki H., Tokunaga A.,
RA Tsutsui K.;
RT "Novel splice variants of amphiphsin I are expressed in retina.";
RL FEBS Lett. 519:185-190(2002).
DR EMBL: AF498097; AM44811.1; -.
DR GO: GO:0006897; P: endocytosis; IEA.
DR GO: GO:0007268; P: synaptic transmission; IEA.
DR InterPro: IPR003005; Amphiphsin.
DR InterPro: IPR003017; Amphiphsin_1.
DR InterPro: IPR006632; BAR.
DR InterPro: IPR004148; BAR_dom.
DR Pfam: PF03114; BAR; 1.
DR PRINTS: PR01251; AMPHIPHYSIN.
DR ProDom: PD003208; Amphiphsin_1; 1.
DR SMART: SM00721; BAR; 1.
FT NON TER 385
SQ SEQUENCE 385 AA; 43753 MW; 4B9AB0A3136711AA CRC64;

Query Match 93.7%; Score 89; DB 4; Length 385;
Best Local Similarity 94.4%; Pred. No. 1.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 134 RIAKSRKLVYDSARHH 151

RESULT 10
Q7TOF7 PRELIMINARY; PRT; 686 AA.
ID AC Q7TOF7
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

Query Match 93.7%; Score 89; DB 4; Length 385;
Best Local Similarity 94.4%; Pred. No. 1.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 134 RIAKSRKLVYDSARHH 151

```

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC054718; AHS4718.1; -.
KW Hypothetical protein.
SQ SEQUENCE 686 AA; 75013 MW; D292E24653A442A5 CRC64;

Query Match 93.7%; Score 89; DB 11; Length 686;
Best Local Similarity 94.4%; Pred. No. 3.4e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 134 RIAKSRKLVYDSARHH 151

```

```

RESULT 11
Q8N4G0 PRELIMINARY; PRT; 695 AA.
ID AC Q8N4G0
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Amphiphsin (Stiff-Man syndrome with breast cancer 128kD
DE autoantigen).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Brain, and Lung;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL: BC034376; AAH34376.1; -.
DR GO: GO:0006897; P: endocytosis; IEA.
DR GO: GO:0007268; P: synaptic transmission; IEA.
DR InterPro: IPR003005; Amphiphsin.
DR InterPro: IPR003017; Amphiphsin_1.
DR InterPro: IPR006632; BAR.
DR InterPro: IPR004148; BAR_dom.
DR InterPro: IPR001452; SH3_dom.
DR Pfam: PF03114; BAR; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR01251; AMPHIPHYSIN.
DR PRINTS: PD00452; SH3DOMAIN.
DR ProDom: PD003208; Amphiphsin_1; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00721; BAR; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 695 AA; 76229 MW; 4FC92EE8B028BBFE CRC64;

Query Match 93.7%; Score 89; DB 4; Length 695;
Best Local Similarity 94.4%; Pred. No. 3.4e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18
DB 134 RIAKSRKLVYDSARHH 151

```

```

RESULT 12
Q21004 PRELIMINARY; PRT; 461 AA.
AC Q21004;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F58G6.1 protein.
GN F58G6.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RX MEDLINE=99069613; PubMed=9851916;
RP SEQUENCE FROM N.A.
RA "Genome sequence of the nematode C.elegans: A platform for
investigating biology";
RL Science 282:2012-2018(1998).
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; Z68217; CAA92465.1; -
DR PIR; T22946; T22946.
DR WormPep; F58G6.1; CE03428.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR004148; BAR.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 461 AA; 51910 MW; 771C89E21EF7BEC8 CRC64;

Query Match 72.6%; Score 69; DB 5; Length 461;
Best Local Similarity 76.5%; Pred. No. 0.0049;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RIAGRGKGLVDYDSARH 17
Db 130 KIEKRGKGLVDYDSAKN 146
: |||||
: |||||

RESULT 13
Q9Y092 PRELIMINARY; PRT; 602 AA.
AC Q9Y092;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Amphiphysin protein (LD19810P).
GN AMPH OR CG8604.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.,

```

```

Anatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Ahril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Terrier S., Fleischmann W.,
Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mout R.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20076404; PubMed=10607911;
RX Razaq A., Su Y., Mehren J.E., Mizuguchi K., Jackson A.P., Gay N.J.,
O'Kane C.J.;
RA "Characterization of the gene for Drosophila amphiphysin.";
RL Gene 241:167-174(2000).
[3]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Zhang Y.Q., Xue H., Broadie K.S.;
RT "Dissecting the functions of amphiphysin.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkeley;
RC Stapleton M., Brokstein P., Horg L., Agbayani A., Carlson J.,
Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Murgall C.J.,
Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AE003821; AAF58486.1; -
DR EMBL; AF171225; AAD46684.1; -
DR EMBL; AJ242855; CAB45188.1; -
DR EMBL; AY061278; AAL28826.1; -
DR HSP; P27986; IPT.
DR FlyBase; FBgn0027356; Amph.
DR GO; GO:0005886; C: plasma membrane; IDA.
DR GO; GO:0045202; C: synaptic junction; IDA.
DR GO; GO:0007269; P: neurotransmitter secretion; NAS.
DR GO; GO:0008104; P: protein localization; IMP.
DR GO; GO:0006937; P: regulation of muscle contraction; IMP.
DR GO; GO:0045313; P: intracellular membrane biogenesis; IMP.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.

```



```

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE VNG1031C.
GN VNG1031C.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Sbrocna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005037; AAG19442.1; -.
DR PIR; F84259; F84259.
KW Complete proteome.
SQ SEQUENCE 169 AA; 17767 MW; 4F280D721091DC8B CRC64;

Query Match 50.5%; Score 48; DB 17; Length 169;
Best Local Similarity 56.2%; Pred. No. 5.5;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 AKGRKLVYDSARHH 18
Db 44 ADRGSKIVDVAVRDH 59

RESULT 18
ID O64842 PRELIMINARY; PRT; 424 AA.
AC O64842;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN A2G26290.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.R.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Freus D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC004484; AAC14522.1; -.
DR PIR; F84658; F84658.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.

```

```

DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0005468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS0108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 424 AA; 48086 MW; C11DD637020B8307 CRC64;

Query Match 49.5%; Score 47; DB 10; Length 424;
Best Local Similarity 53.3%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 KEGRKLVYDSARHH 18
Db 382 KEGKFFVDINKFRHH 396

RESULT 19
Q39229 PRELIMINARY; PRT; 425 AA.
ID Q39229;
AC Q39229;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine protein kinase.
DE ARSK1.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Columbia;
RX MEDLINE=95384198; PubMed=7655506;
RA Hwang I., Goodman H.M.;
RT "An Arabidopsis thaliana root-specific kinase homolog is induced by
dehydration, ABA, and NaCl."
RL Plant J. 8:37-43(1995).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; L22302; AA81538.1; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0005468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS0108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 425 AA; 48311 MW; 77998E7606EC94E CRC64;

Query Match 49.5%; Score 47; DB 10; Length 425;
Best Local Similarity 53.3%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 KEGRKLVYDSARHH 18
Db 383 KEGKFFVDINKFRHH 397

```

## RESULT 20

Q9LEA1 ID Q9LEA1 PRELIMINARY; PRT; 441 AA.  
 AC Q9LEA1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Asparagine synthetase (ASN3) (Fragment) (Fragment).  
 GN T31P16 230.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Kalicki J.,  
 RA Wohldmann P., Smith A., Bancroft I., Mewes H.W., Rudd S., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AL356332; CAB92065.1; -.  
 DR HSSP; P22106; 1CT9.  
 DR GO; GO:0004066; F:asparagine synthase (glutamine-hydrolyzing)...; IEA.  
 DR GO; GO:0006529; P:asparagine biosynthesis; IEA.  
 DR InterPro; IPR001962; Asn\_synthase.  
 DR Pfam; PF00733; Asn\_synthase; 1.  
 FT NON\_TER  
 SQ SEQUENCE 441 AA; 49619 MW; 7C0B9FCD15D536EA CRC64;

Query Match 49.5%; Score 47; DB 10; Length 441;  
 Best Local Similarity 53.3%; Pred. No. 23;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KGRKLVYDYSARHH 18  
 DB 139 KAGREVADYLGTRHH 153

## RESULT 21

Q9ZST7 ID Q9ZST7 PRELIMINARY; PRT; 578 AA.  
 AC Q9ZST7;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Asparagine synthetase.  
 GN ASN3  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99097830; PubMed=9881155;  
 RA Lam H.M., Hsieh M.H., Coruzzi G.;  
 RL "Reciprocal regulation of distinct asparagine synthetase genes by  
 light and metabolites in Arabidopsis thaliana.";  
 RL Plant J. 16:345-353(1998).  
 DR EMBL; AF095452; AAC72836.1; -.  
 DR FIR; T51888; T51888.  
 DR HSSP; P22106; 1CT9.  
 DR GO; GO:0004066; F:asparagine synthase (glutamine-hydrolyzing)...; IEA.  
 DR GO; GO:0006529; P:asparagine biosynthesis; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001962; Asn\_synthase.  
 DR InterPro; IPR006426; Asn\_synthase.  
 DR InterPro; IPR000583; GATase\_2.  
 DR Pfam; PF00733; Asn\_synthase; 1.

DR Pfam; PF00310; GATase\_2; 1.  
 DR TIGRfams; TIGR01536; asn\_synth\_AEB; 1.  
 DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
 SQ SEQUENCE 578 AA; 65223 MW; 5BBD2FC7A04636BD CRC64;

Query Match 49.5%; Score 47; DB 10; Length 578;  
 Best Local Similarity 53.3%; Pred. No. 31;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KGRKLVYDYSARHH 18  
 DB 276 KAGREVADYLGTRHH 290

## RESULT 22

Q9ZST6 ID Q9ZST6 PRELIMINARY; PRT; 578 AA.  
 AC Q9ZST6;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Asparagine synthetase.  
 GN ASN2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99097830; PubMed=9881155;  
 RA Lam H.M., Hsieh M.H., Coruzzi G.;  
 RL "Reciprocal regulation of distinct asparagine synthetase genes by  
 light and metabolites in Arabidopsis thaliana.";  
 RL Plant J. 16:345-353(1998).  
 DR EMBL; AF095453; AAC72837.1; -.  
 DR HSSP; P22106; 1CT9.

DR GO; GO:0004066; F:asparagine synthase (glutamine-hydrolyzing)...; IEA.  
 DR GO; GO:0006529; P:asparagine biosynthesis; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001962; Asn\_synthase.  
 DR InterPro; IPR006426; Asn\_synth\_AEB.  
 DR InterPro; IPR000583; GATase\_2.  
 DR Pfam; PF00733; Asn\_synthase; 1.  
 DR Pfam; PF00310; GATase\_2; 1.  
 DR TIGRfams; TIGR01536; asn\_synth\_AEB; 1.  
 DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
 SQ SEQUENCE 578 AA; 65077 MW; D05B73C63BF3B663 CRC64;

Query Match 49.5%; Score 47; DB 10; Length 578;  
 Best Local Similarity 53.3%; Pred. No. 31;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KGRKLVYDYSARHH 18  
 DB 276 KAGREVADYLGTRHH 290

## RESULT 23

Q9LV77 ID Q9LV77 PRELIMINARY; PRT; 578 AA.  
 AC Q9LV77;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Asparagine synthetase (AT5G65010/MXK3 25).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.





Homo sapiens (Human) . Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OK NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nunes D.N., Dias-Neto E., Brentani R.R., Camargo A.A.;  
 RT "Cloning and characterization of two human homologs of the enhancer of  
 RT polycomb gene (EPC1) from *Drosophila*."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF286905; AAK60501.1; -.  
 SQ SEQUENCE 763 AA; 85219 MW; 588D375F40151D03 CRC64;  
  
 Query Match 47.9%; Score 45.5; DB 4; Length 763;  
 Best Local Similarity 43.5%; Pred. No. 74;  
 Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;  
  
 QY 1 RIAKRGKLV-----DYDSARHH 18  
 : : : : :  
 Db 408 RVGRGRLVLLDRAHSDYDSVFHH 430  
 : : : : :  
  
 RESULT 30  
 Q8NE21 PRELIMINARY; PRT; 813 AA.  
 ID Q8NE21  
 AC Q8NE21  
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DE Similar to enhancer of polycomb 1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Testis;  
 RA Strausberg R.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC036529; AAK36529.1; -.  
 SQ SEQUENCE 813 AA; 90893 MW; 1A7DE9D14C291D11 CRC64;  
  
 Query Match 47.9%; Score 45.5; DB 4; Length 813;  
 Best Local Similarity 43.5%; Pred. No. 79;  
 Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;  
  
 QY 1 RIAKRGKLV-----DYDSARHH 18  
 : : : : :  
 Db 458 RVGRGRLVLLDRAHSDYDSVFHH 480  
 : : : : :  
  
 RESULT 31  
 Q9H2F5 PRELIMINARY; PRT; 836 AA.  
 ID Q9H2F5  
 AC Q9H2F5  
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DE Enhancer of polycomb.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20564312; PubMed=10976108;  
 RA Shimono Y., Murakami H., Hasegawa Y., Takahashi M.;  
 RT "RET Finger Protein is a Transcriptional Repressor and Interacts with  
 RT Enhancer of Polycomb That Has Dual Transcriptional Functions."  
 RL J. Biol. Chem. 275:39411-39419 (2000).  
 DR EMBL: AF277374; AAG41402.1; -.  
 DR Genbank: HGNC:19876; EPC1.  
 DR GO: GO:0005634; C:nucleus; IDA.  
 DR GO: GO:0016563; F:transcriptional activator activity; TAS.  
 DR GO: GO:0016564; F:transcriptional repressor activity; IDA.  
 DR GO: GO:0006350; P:transcription; TAS.

SQ SEQUENCE 836 AA; 93463 MW; E9E89699E73336B5 CRC64;  
  
 Query Match 47.9%; Score 45.5; DB 4; Length 836;  
 Best Local Similarity 43.5%; Pred. No. 81;  
 Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;  
  
 QY 1 RIAKRGKLV-----DYDSARHH 18  
 : : : : :  
 Db 458 RVGRGRLVLLDRAHSDYDSVFHH 480  
 : : : : :  
  
 RESULT 32  
 Q852V9 PRELIMINARY; PRT; 128 AA.  
 ID Q852V9  
 AC Q852V9  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DE Similar to enhancer of polycomb 1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=205877;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22592660; PubMed=12705866;  
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,  
 RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,  
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
 RA Hatfull G.F.;  
 RT "Origins of highly mosaic mycobacteriophage genomes."  
 RL Cell 113:171-182 (2003).  
 DR EMBL: AY129337; AAN16878.1; -.  
 SQ SEQUENCE 128 AA; 14386 MW; 1A48FD39E58CBA70 CRC64;  
  
 Query Match 47.4%; Score 45; DB 9; Length 128;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 6 GRKLVYDSARH 17  
 : : : : :  
 Db 101 GRKLVYDSARH 112  
 : : : : :  
  
 RESULT 33  
 Q9UOX9 PRELIMINARY; PRT; 160 AA.  
 ID Q9UOX9  
 AC Q9UOX9  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
 DE Hypothetical protein.  
 GN L171.04.  
 OS Leishmania major.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OC NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Friedlin;  
 RA Tosato V., Bruschi C.V., Ivens A.C., Murphy L., Quail M.,  
 RA Rajandream M.A., Barrell B.G.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Friedlin;  
 RX MEDLINE=98146435; PubMed=9477341;  
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 RA Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome."  
 RL Genome Res. 8:135-145 (1998).  
 DR EMBL: AL133435; CAB62807.1; -.  
 KW Hypothetical protein.

OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;	
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=CS7BL/6; TISSUE=Brain;
RC	Strauberg R.;
RA	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL	EMBL; BC043061; AAH43061.1; -
DR	InterPro; IPR001440; TPR.
DR	InterPro; IPR008941; TPR-like.
DR	Pfam; PF00515; TPR; 3.
DR	SMART; SMO0028; TPR; 3.
DR	SEQUENCE 171 AA; 19998 MW; 6BD8AF10567F7B1D CRC64;
SQL	
Query Match	47.4%; Score 45; DB 11; Length 171;
Best Local Similarity	80.0%; Pred. No. 18;
Matches 8; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Qy	9 LVVDYSARHH 18
Db	121 LQYDRARHH 130
RESULT 36	
Q9ZXLO	
ID	Q9ZXLO PRELIMINARY; PRT; 190 AA.
AC	Q9ZXLO;
DT	01-MAY-1999 (TEMBLrel. 10, Created)
DT	01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT	01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE	Orf16.
OS	Bacteriophage phi CTX.
OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
NCBI_TaxID=35343;	
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=phiCTX-c;
RC	Medline=99157549; PubMed=10027959;
RA	Nakayama K., Kanaya S., Ohnishi M., Terawaki Y., Hayashi T.;
RL	"The complete nucleotide sequence of phiCTX, a cytotoxin-converting phage of Pseudomonas aeruginosa: implications for phage evolution and horizontal gene transfer via bacteriophage.";
DR	Mol. Microbiol. 31:399-419(1999).
[2]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=phiCTX-c;
RC	Medline=90014160; PubMed=2507866;
RA	Hayashi T., Kamio Y., Hishinuma F., Usami Y., Titani K., Terawaki Y.;
RL	"Pseudomonas aeruginosa cytotoxin; the nucleotide sequence of the gene and the mechanism of activation of the protoxin.";
DR	Mol. Microbiol. 3:861-868(1989).
DR	EMBL; AB008550; BAA36243.1; -
DR	InterPro; IPR006531; Phage.P2.V.
DR	Pfam; PF04717; phage_base.V.1.
DR	TIGRFAMs; TIGR01644; Phage.P2.V.1.
SQL	SEQUENCE 190 AA; 19961 MW; 8C8FE335568FDDCA CRC64;
Query Match	47.4%; Score 45; DB 9; Length 190;
Best Local Similarity	44.4%; Pred. No. 20;
Matches 8; Conservative	3; Mismatches 7; Indels 0; Gaps 0;
Qy	1 RIAKRGKLVYDSARHH 18
Db	98 RTYRGAVIEYDSARHH 115
RESULT 37	

```

Q8KD94
ID O8KD94 PRELIMINARY; PRT; 524 AA.
AC O8KD94
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Type III restriction system endonuclease, putative.
GN CT1160.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=1203901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Niernan W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AB012877; AA072393.1; -.
DR TIGR; CT1160; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR002052; N6_Mtase.
DR PRINTS; PR00507; N12N6MTFRASE.
DR PROSITE; PS00092; N6_MTASE; 1.
DR ENDONUCLEASE; Complete proteome.
KW Endonuclease.
SQ SEQUENCE 524 AA; 59223 MW; A3F705751D3ABAA7 CRC64;

Query Match 47.4%; Score 45; DB 16; Length 524;
Best Local Similarity 49.2%; Pred. No. 60;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 KRGRKLVYDSAR 16
DB 266 KRIRKLVYENQA 278

RESULT 38
Q7WDT3 PRELIMINARY; PRT; 263 AA.
ID Q7WDT3
AC Q7WDT3
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Arac family regulatory protein.
GN B4904.
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640436; CAE39597.1; -.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 263 AA; 29173 MW; 143B129BB0E436B0 CRC64;

Query Match 46.3%; Score 44; DB 16; Length 263;
Best Local Similarity 46.7%; Pred. No. 42;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGRKLVYDSAR 18
DB 8 RAGHEVVDYDHPGHH 22

RESULT 40
Q7VSK1 PRELIMINARY; PRT; 297 AA.
ID Q7VSK1
AC Q7VSK1
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Arac family regulatory protein.
GN B3798.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.

```

```

OX NCBI_TaxID=520;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Tohana I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=2287954; PubMed=12310271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger C.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Urwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL: BX640422; CAE44053.1; -.
KW Complete proteome.
SQ SEQUENCE 297 AA; 32461 MW; 5A462C5DFC948D23 CRC64;

Query Match 46.3%; Score 44; DB 16; Length 297;
Best Local Similarity 46.7%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 KRGKRLVDYDSARH 18
Db : : : : :
42 RAGHEVVDYDHPGHH 56

RESULT 41
Q8GDV2 PRELIMINARY; PRT; 524 AA.
ID Q8GDV2
AC Q8GDV2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Cyclic acid synthase (fragment).
OS Helicobacter mobilis.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;
OC Helicobacter.
OX NCBI_TaxID=28064;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=22337798; PubMed=12446909;
RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
RA Blankenship R.E.;
RT "Whole-genome analysis of photosynthetic prokaryotes.";
RL Science 298:1616-1620(2002).
RN [2]
RN SEQUENCE FROM N.A.
RA Liolios K.G., Chu L., Ostrovskaya O., Mendenbaeva N., Koukharenko V.,
RA Gerdes S., Kyripides N., Overbeek R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY142870; RAN87474.1; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0042242; P: coenzyme A, C-di- and C-di- amide synthase activity; IEA.
DR GO; GO:0003236; P: vitamin B12 biosynthesis; IEA.
DR InterPro: IPR002596; CblA_P.
DR InterPro: IPR004459; COBQ_synth.
DR Pfam: PF01656; CblA; 1.
DR TIGRFAMs: TIGR00313; cobQ; 1.
FT NON_TER 524
SQ SEQUENCE 524 AA; 56649 MW; 52E9B5770880660A CRC64;

Query Match 46.3%; Score 44; DB 2; Length 524;
Best Local Similarity 30.0%; Pred. No. 88;
Matches 12; Conservative 2; Mismatches 4; Indels 22; Gaps 1;

QY 1 RIAGKRGKLVYD-----DSARH 18
Db : : : : :
418 RIRGGEELVDIVDADGVIGSYLHGFDNDSLHH 457

```

```

RESULT 42
Q8ZN24 PRELIMINARY; PRT; 282 AA.
ID Q8ZN24
AC Q8ZN24;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative ABC superfamily (Membrane) transport protein.
DE YFHH OR STM2572.
GN Salmonella typhimurium.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking I., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL: AE008817; AAL21466.1; -.
DR GO; GO:0005529; F: sugar binding; IEA.
DR GO; GO:0003700; F: transcription factor activity; IEA.
DR GO; GO:0005979; P: carbohydrate metabolism; IEA.
DR GO; GO:0006355; P: regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR00281; HTH_FpIR.
DR InterPro: IPR001347; SIS-.
DR Pfam: PF01418; HTH_6; 1.
DR Pfam: PF01380; SIS; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 282 AA; 30925 MW; 170726A1AF8A25 CRC64;

Query Match 45.8%; Score 43.5; DB 16; Length 282;
Best Local Similarity 47.6%; Pred. No. 54;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 2 IAXGRKLVYD-----DSARH 17
Db : : : : :
14 LAQSDRLADYLLAQPDTH 34

RESULT 43
Q824L1 PRELIMINARY; PRT; 282 AA.
ID Q824L1
AC Q824L1;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
DE YFHH OR STY2818 OR T0285.
GN Salmonella typhi.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moulé S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;

```

Qy 2 IAKRGRKLVYDS 14

DR MGD; MGI:2441685; Ankrd5.  
DR GO; GO:0005509; F:calcium ion binding; IEA.



DR InterPro; IPR003130; GED.  
DR InterPro; IPR001849; PH.  
DR Pfam; PF00350; dynamin\_1.  
DR Pfam; PF01031; dynamin\_2; 1.  
DR Pfam; PF02212; GED; 1.  
DR Pfam; PF00169; PH; 1.  
DR PRINTS; PR00195; DYNAMIN.  
DR SMART; SMO0053; DYNG; 1.  
DR SMART; SMO0302; GED; 1.  
DR SMART; SMO0233; PH; 1.  
DR PROSITE; PS00410; DYNAMIN; 1.  
DR PROSITE; PS00003; PH\_DOMAIN; 1.  
SQ SEQUENCE 863 AA; 97190 MW; 8BDBB091CB10D264 CRC64;

Query Match 45.3%; Score 43; DB 11; Length 863;  
Best Local Similarity 46.2%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 RGRKLVYDSARH 17  
:|:|:|:|:  
Db 87 KGKFTDFEVRH 99

## RESULT 50

Q9QXL9 PRELIMINARY; PRT; 869 AA.  
AC Q9QXL9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dynamlin il1bb isoform.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=96355732; PubMed=9752097;  
RA Cook T., Mesa K., Urrutia R.;  
RT "Three dynamin-encoding genes are differentially expressed in  
developing rat brain."  
RL J. Neurochem. 67:927-931(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Cook T., Mesa K., Urrutia R.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.  
DR EMBL; AF201839; AAF07848.1; -.  
DR HSSP; Q05193; LDYN.  
DR GO; GO:0003924; F:GTPase activity; IEA.  
DR InterPro; IPR001401; Dynamlin.  
DR InterPro; IPR000375; Dynamlin\_central.  
DR InterPro; IPR003130; GED.  
DR InterPro; IPR001849; PH.  
DR Pfam; PF00350; dynamin; 1.  
DR Pfam; PF01031; dynamin\_2; 1.  
DR Pfam; PF02212; GED; 1.  
DR Pfam; PF00169; PH; 1.  
DR PRINTS; PR00195; DYNAMIN.  
DR SMART; SMO0053; DYNG; 1.  
DR SMART; SMO0302; GED; 1.  
DR SMART; SMO0233; PH; 1.  
DR PROSITE; PS00410; DYNAMIN; 1.  
DR PROSITE; PS00003; PH\_DOMAIN; 1.  
SQ SEQUENCE 869 AA; 97914 MW; 00B41E41E5425EAD CRC64;

Query Match 45.3%; Score 43; DB 11; Length 869;  
Best Local Similarity 46.2%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 RGRKLVYDSARH 17

Db 87 KGKFTDFEVRH 99  
:|:|:|:|:

Search completed: March 4, 2004, 17:46:25  
Job time : 49.3871 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:19:24 ; Search time 28.4344 Seconds  
(without alignments)  
2196.038 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_1\_221

Perfect score: 1125  
Sequence: 1 MAEGKAGGAGLFAKQVKK.....IGCVTFQNIENLRDVFYR 221

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1125	100.0	564	AAB62100	Human bri
2	1114.5	99.1	414	ABG13005	Novel hum
3	1111.5	98.8	594	ABG20887	Novel hum
4	1111.5	98.8	594	ABG20324	Novel hum
5	1104.5	98.2	594	ABG22457	Novel hum
6	1091.5	97.0	572	ABG21411	Novel hum
7	833.5	74.1	960	ABG20886	Novel hum
8	833.5	74.1	960	ABG20323	Novel hum
9	768.5	68.3	425	ABG22456	Novel hum
10	688	61.2	434	AABW05391	Mouse SH3
11	683	60.7	252	AAB62101	Bar domain
12	683	60.7	482	ABU89771	Protein d
13	671.5	59.7	451	AABW06602	Human box
14	671.5	59.7	451	AABW47295	Human Bin
15	671.5	59.7	451	AAB36881	Human Bin
16	671.5	59.7	451	AABW4504	Human Bin
17	663.5	59.0	588	ADP45951	Rat Prote
18	663.5	59.0	588	ADP360848	Rat Prote
19	663.5	59.0	588	ADP360840	Rat Prote
20	663.5	59.0	588	ADP360844	Rat Prote
21	663.5	59.0	588	ADP360852	Rat Prote
22	657.5	58.4	593	ADP360842	Human Pro
23	657.5	58.4	593	ADP45953	Human Pro
24	657.5	58.4	593	ADP60850	Human Pro
25	657.5	58.4	593	ADP60846	Human Pro

## ALIGNMENTS

### RESULT 1

AAB62100  
ID AAB62100 standard; protein; 564 AA.  
XX  
AC AAB62100;  
XX  
DT 29-MAY-2001 (first entry)  
XX  
DE Human bridging integrator-2 (Bin2) protein.  
XX  
KW Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
KW Hyperplastic disease; cytostatic; cell growth regulator;  
KW chromosome 4q22.1.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Peptide  
FT Peptide  
FT Peptide  
FT Peptide  
XX  
PN WO200116158-A2.  
XX  
PD 08-MAR-2001.  
XX  
PF 30-AUG-2000; 2000WO-US023723.  
XX  
PR 31-AUG-1999; 99US-0151554P.  
XX  
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
XX  
PI Prendergast GC, Ge K;  
XX  
DR WPI; 2001-235087/24.  
XX  
DR N-PSDB; AAF57268.  
XX  
PT New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
PT useful for regulating cell growth, and for diagnosing or treating  
PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
or hepatocarcinoma.  
XX  
PS Claim 1; Fig 1A-C; 62pp; English.  
XX  
CC This represents a human bridging integrator-2 (Bin2) protein. The Bin2  
CC protein can be expressed by standard recombinant methodology. The Bin2  
CC proteins or peptides are useful in regulating cell growth, cell survival,  
CC differentiation, endocytosis and actin organization. These peptides or

26	657.5	58.4	593	7	ADE60854	Human Pro
27	625.5	55.6	130	4	ABG22449	Novel hum
28	625.5	55.6	130	4	ABG20881	Novel hum
29	625.5	55.6	130	4	ABG21407	Novel hum
30	625.5	55.6	130	4	ABG20318	Novel hum
31	620.5	55.2	683	7	ADD44887	Rat Prote
32	620.5	55.2	695	5	AU75110	Synaptic
33	620.5	55.2	695	5	ADP44889	Human Pro
34	619.5	55.1	695	5	AAE22091	Human amp
35	605	53.8	453	6	ABR69631	Human SH3
36	559.5	49.7	404	2	AAW05392	Human NS
37	527	46.8	108	5	ABB06161	Human NS
38	527	46.8	137	4	ABG21410	Novel hum
39	429.5	38.2	602	4	ABG63948	Drosophil
40	224	19.9	300	4	ABG20320	Novel hum
41	224	19.9	300	4	ABG22452	Novel hum
42	190.5	16.9	253	4	ABG20883	Novel hum
43	190.5	16.9	253	4	ABG20319	Novel hum
44	190.5	16.9	253	4	ABG22450	Novel hum
45	190.5	16.9	253	4	ABG20882	Novel hum



PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 20; SEQ ID NO 51246; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 594 AA;  
  
Query Match 98.8%; Score 1111.5; DB 4; Length 594;  
Best Local Similarity 99.1%; Pred. No. 4.3e-84;  
Matches 220; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
  
QY 1 MAEGKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSANFYQQQAEHKL 60  
Db |||||  
QY 30 MAEGKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSANFYQQQAEHKL 89  
Db |||||  
QY 61 YKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEELADQ 120  
Db |||||  
QY 90 YKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEELADQ 149  
Db |||||  
QY 121 AVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAESEFNKA 179  
Db |||||  
QY 150 AVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAESEFNKA 209  
Db |||||  
QY 180 QTVFEDLNQELLELPILYNSRIGCVYTFIFONISNLRDVFYR 221  
Db |||||  
QY 210 QTVFEDLNQELLELPILYNSRIGCVYTFIFONISNLRDVFYR 251  
Db |||||  
  
RESULT 4  
ABG20324  
ID ABG20324 standard; protein; 594 AA.  
XX  
AC ABG20324;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #20315.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX

PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX N-PSDB; AAS84511.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 20; SEQ ID NO 50693; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 594 AA;  
  
Query Match 98.8%; Score 1111.5; DB 4; Length 594;  
Best Local Similarity 99.1%; Pred. No. 4.3e-84;  
Matches 220; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
  
QY 1 MAEGKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSANFYQQQAEHKL 60  
Db |||||  
QY 30 MAEGKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSANFYQQQAEHKL 89  
Db |||||  
QY 61 YKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEELADQ 120  
Db |||||  
QY 90 YKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEELADQ 149  
Db |||||  
QY 121 AVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAESEFNKA 179  
Db |||||  
QY 150 AVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAESEFNKA 209  
Db |||||  
QY 180 QTVFEDLNQELLELPILYNSRIGCVYTFIFONISNLRDVFYR 221  
Db |||||  
QY 210 QTVFEDLNQELLELPILYNSRIGCVYTFIFONISNLRDVFYR 251  
Db |||||  
  
RESULT 5  
ABG22457  
ID ABG22457 standard; protein; 594 AA.  
XX  
AC ABG22457;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #22448.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

PF 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

XX N-PSDB; AAS85644.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 52816; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 594 AA;

Query Match 98.2%; Score 1104.5; DB 4; Length 594;

Best Local Similarity 98.6%; Pred. No. 1.6e-83;

Matches 219; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MAEGKAGGAGLFAKQVKFSAQKVLQKLGKAVETKDERPQSASNFYQQQAEHKL 60

Db 30 MAEGKAGGAGLFAKQVKFSAQKVLQKLGKAVETKDERPQSASNFYQQQAEHKL 89

QY 61 YKDLKNFLSAVKVMHESSEKRVSETLQEIYSSWGDGHEELKAIWVNNLLWEDYEKLADQ 120

Db 90 YKDLKNFLSAVKVMHESSEKRVSETLQEIYSSWGDGHEELKAIWVNNLLWEDYEKLADQ 149

QY 121 AVRTMEIYVAQFEIKERIAKRGKLVDSARHLEAVQNA-KDEAKTAKAEFEFNA 179

Db 150 AVRTMEIYVAQFEIKERIAKRGKLVDSARHLEAVQNA-KDEAKTAKAEFEFNA 209

QY 180 QTFPEDLNQELLELPILYNSRIGCVYTFQNISNLRDVFYR 221

Db 210 QTFPEDLNQELLELPILYNSRIGCVYTFQNISNLRDVFYR 251

RESULT 6

ABG21411

ID ABG21411 standard; protein; 572 AA.

XX AC ABG21411;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #21402.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS85598.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 51770; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 572 AA;

Query Match 97.0%; Score 1091.5; DB 4; Length 572;

Best Local Similarity 98.2%; Pred. No. 1.9e-82;

Matches 218; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MAEGKAGGAGLFAKQVKFSAQKVLQKLGKAVETKDERPQSASNFYQQQAEHKL 60

Db 8 MAEGKAGGAGLFAKQVKFSAQKVLQKLGKAVETKDERPQSASNFYQQQAEHKL 67

QY 61 YKDLKNFLSAVKVMHESSEKRVSETLQEIYSSWGDGHEELKAIWVNNLLWEDYEKLADQ 120

Db 68 YKDLKNFLSAVKVMHESSEKRVSETLQEIYSSWGDGHEELKAIWVNNLLWEDYEKLADQ 127

QY 121 AVRTMEIYVQAFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 179  
 DB 128 AVRTMEIYVQAFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKA 187  
 QY 180 QTVFEDLNQELLELPILYNSRIGCYVTIFQNTSNLRDVFYR 221  
 DB 188 QTVFEDLNQELLELPILYNSRIGCYVTIFQNTSNLRDVFYR 229

## RESULT 7

ABG20886  
 ID ABG20886 standard; protein; 960 AA.

XX AC ABG20886;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #20877.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS85073.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity.

XX PS Claim 20; SEQ ID NO 51245; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 sequences. (I) is useful as hybridisation probes, polymerase chain  
 reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 and in recombinant production of (II). The polynucleotides are also used  
 in diagnostics as expressed sequence tags for identifying expressed  
 genes. (I) is useful in gene therapy techniques to restore normal  
 activity of (II) or to treat disease states involving (II). (II) is  
 useful for generating antibodies against it, detecting or quantitating a  
 polypeptide in tissue, as molecular weight markers and as a food  
 supplement. (II) and its binding partners are useful in medical imaging  
 of sites expressing (II). (I) and (II) are useful for treating disorders  
 involving aberrant protein expression or biological activity. The  
 polypeptide and polynucleotide sequences have applications in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits to assess biodiversity  
 and to produce other types of data and products dependent on DNA and  
 amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 amino acid sequences of the invention. Note: The sequence data for this  
 patent did not appear in the printed specification, but was obtained in  
 electronic format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 960 AA;

Query Match

Best Local Similarity 74.1%; Score 833.5; DB 4; Length 960;  
 Pred. No. 1.2e-60;

Matches 168; Conservative 3; Mismatches 5; Indels 3; Gaps 2;  
 QY 46 SASNF--YQQAEGHKLKXDLNPLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIV 103  
 DB 144 SAFHFFIYKCRAGHKLKXDLNPLSAVKVMHSSKRVSETLQFIYSEWYGHLEELKAIV 203  
 QY 104 WNDLLWEDYEKLDQAVRTMEIYVQAFSEIKERIAKRGKLVYDSARHLEAVQNA- 162  
 DB 204 WNDLLWEDYEKLDQAVRTMEIYVQAFSEIKERIAKRGKLVYDSARHLEAVQNA 263  
 QY 163 KDEAKTAKAEFEFNKAQTVFEDLNQELLELPILYNSRIGCYVTIFQNTSNLRDVFYR 221  
 DB 264 KDEAKTAKAEFEFNKAQTVFEDLNQELLELPILYNSRIGCYVTIFQNTSNLRDVFYR 322

## RESULT 8

ABG20323

ID ABG20323 standard; protein; 960 AA.

XX AC ABG20323;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #20314.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS84510.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity.

XX PS Claim 20; SEQ ID NO 50682; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 sequences. (I) is useful as hybridisation probes, polymerase chain  
 reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 and in recombinant production of (II). The polynucleotides are also used  
 in diagnostics as expressed sequence tags for identifying expressed  
 genes. (I) is useful in gene therapy techniques to restore normal  
 activity of (II) or to treat disease states involving (II). (II) is  
 useful for generating antibodies against it, detecting or quantitating a  
 polypeptide in tissue, as molecular weight markers and as a food  
 supplement. (II) and its binding partners are useful in medical imaging  
 of sites expressing (II). (I) and (II) are useful for treating disorders  
 involving aberrant protein expression or biological activity. The  
 polypeptide and polynucleotide sequences have applications in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits to assess biodiversity  
 and to produce other types of data and products dependent on DNA and  
 amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 amino acid sequences of the invention. Note: The sequence data for this  
 patent did not appear in the printed specification, but was obtained in  
 electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 960 AA;

SQ

Query Match 74.1%; Score 833.5; DB 4; Length 960;

Best Local Similarity 93.9%; Pred. No. 1.2e-60;

Matches 168; Conservative 3; Mismatches 5; Indels 3; Gaps 2;

QY 46 SASNP--YQQAEGHKLKDLKNFLSAVKVMHSSKVSSETLQEIYSSWDGHEELKATV 103

DB 144 SAFFHYKCAEGHKLKDLKNFLSAVKVMHSSKVSSETLQEIYSSWDGHEELKATV 203

QY 104 WNDLLMEDYBEKLAQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA- 162

DB 204 WNDLLMEDYBEKLAQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA- 263

QY 163 KQDAKAKAESEFNKAQTVFEDLNQELLELPILYNSRIGCVTIFQNISNLRDVFYR 221

DB 264 KQDAKAKAESEFNKAQTVFEDLNQELLELPILYNSRIGCVTIFQNISNLRDVFYR 322

#### RESULT 9

ABG22456

ID ABG22456 standard; protein; 425 AA.

XX

AC ABG22456;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #22447.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX

XX (HYSE-) HYSEQ INC.

PA

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

XX

DR N-PSDB; RAS86643.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX

PS Claim 20; SEQ ID NO 52815; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABC0010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 425 AA;

Query Match 68.3%; Score 768.5; DB 4; Length 425;

Best Local Similarity 90.1%; Pred. No. 1.1e-55;

Matches 155; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

QY 51 YQQAEGHKLKDLKNFLSAVKVMHSSKVSSETLQEIYSSWDGHEELKATVWNNDLLW 110

DB 10 FTQAEGHKLKDLKNFLSAVKVMHSSKVSSETLQEIYSSWDGHEELKATVWNNDLLW 69

QY 111 EDYEEKLAQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KXDEAKT 169

DB 70 EDYEEKLAQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KXDEAKT 129

QY 170 AKAESEFNKAQTVFEDLNQELLELPILYNSRIGCVTIFQNISNLRDVFYR 221

DB 130 AKAESEFNKAQTVFEDLNQELLELPILYNSRIGCVTIFQNISNLRDVFYR 181

#### RESULT 10

AAW05391

ID AAW05391 standard; protein; 434 AA.

XX

AC AAW05391;

XX

DT 18-FEB-1998 (first entry)

XX

DE Mouse SH3p9 protein.

XX

KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KW cellular signalling element; cellular structural element; malignancy;

KW protein identification; functional domain; protein screening;

KW cellular signal transduction process.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

XX

FT Misc-difference 433

XX

XX /note= "encoded by CTA"

XX

XX WO9631625-A1.

XX

PD 10-OCT-1996.

XX

XX 04-APR-1996; 96WO-US004454.

XX

XX 07-APR-1995; 95US-00417872.

XX

XX 03-APR-1996; 96US-00630915.

XX

XX (CYTO-). CYTOGEN CORP.

XX

XX (YUNC-) UNIV NORTH CAROLINA.

XX

XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;

XX

XX WPI; 1996-465045/46.

XX

XX N-PSDB; AAT39791.

XX

XX Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

XX Claim 54; Fig 35; 174pp; English.

XX AAW05386-W05403 represent novel human and mouse Src-homology region 3

CC (SH3) domain containing proteins that can be used in the method of the

CC invention. SH3 domain containing proteins play a role in signalling and  
 CC structural elements of cells. The method of the invention is for  
 CC identifying polypeptides containing functional domains of interest  
 CC (especially SH3 domains). The method comprises contacting a multivalent  
 CC recognition unit (RU) complex with a number of peptides and identifying  
 CC polypeptides having a selective binding affinity for the RU complex. The  
 CC method is based on functional similarities and does not rely on sequence  
 CC similarities. Prior methods only gave limited success for identifying  
 CC proteins which contain an SH3 domain due to the minimal sequence homology  
 CC among known SH3 proteins. It has been found that small peptide RUS in  
 CC multivalent form have reduced specificity for a given functional domain  
 CC compared to monomer RUS. Multivalent RU complexes are particularly suited  
 CC to screening for polypeptides containing functional domains that are  
 CC similar to, but not identical in sequence to, the original target  
 CC functional domain. The new method enables proteins having a common  
 CC function to be identified. Identification of novel SH3 proteins will be  
 CC useful for a better understanding of cell growth, malignancy, signal  
 CC transduction processes, etc. New candidate drugs can be identified, and  
 CC their specificities (e.g. pharmacological activities) can be assessed  
 CC using the method of the invention  
 XX  
 XX SQ Sequence 434 AA;

Query Match 61.2%; Score 688; DB 2; Length 434;  
 Best Local Similarity 61.9%; Pred. No. 5.7e-49;  
 Matches 138; Conservative 30; Mismatches 53; Indels 2; Gaps 2;  
 QY 1 MAE-GRAGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASNYQQAGHGK 59  
 Db 1 MAEMSGKGVYTAGKIASNVQKLTQRAQEKVLQKLGKADETKDEQFCQVQNFNKLTEGTR 60  
 QY 60 LYKDLNFLSAVKVMHSSKEVSETLQETYSSEWDGHEELKAIWNNDLLWEDYEELKAD 119  
 Db 61 LQKDLRTYLASVKAMHEASKLSECLQEVYEPWPGRDEANKIAENNDLLWMDYHQKLV 120  
 QY 120 QAVRTMEIYVAQSEIKERIAKRGKLVYDSDARHLEAVQNA-KKDEAKTAKAEFEFNK 178  
 Db 121 QALLTMDTYLGGFPDIKSIKRIAKRGKLVYDSDARHLYESLQTKKKDEAKIAKEEBLIK 180  
 QY 179 AQTVPEDLNQELLELPILYNSRIGCVYTFONISNLRDVFYR 221  
 Db 181 AOKVFEEMNVDLQELPSLWNSRVGFYVNTFQSIAGLEENFHK 223

RESULT 11  
 AAB62101  
 ID AAB62101 standard; protein; 252 AA.  
 AC AAB62101;  
 DT 29-MAY-2001 (first entry)  
 XX Bar domain of Bin1 protein.  
 DE Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
 KW hyperplastic disease; cytostatic; cell growth regulator; Bin1;  
 KW chromosome 4q22.1.  
 XX Homo sapiens.  
 OS WO200116158-A2.  
 PN 08-MAR-2001.  
 PD 30-AUG-2000; 2000WO-US023723.  
 PF 31-AUG-1999; 99US-0151554P.  
 PR (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
 PA Prendergast GC, Ge K;  
 PI WPI; 2001-235087/24.  
 DR

XX  
 PT New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
 PT useful for regulating cell growth, and for diagnosing or treating  
 PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
 PT or hepatocarcinoma.  
 XX  
 PS Example 3; Fig 2; 62pp; English.  
 XX  
 CC The invention provides a human bridging integrator-2 (Bin2) protein. The  
 CC protein can be expressed by standard recombinant methodology. The Bin2  
 CC proteins or peptides are useful in regulating cell growth, cell survival,  
 CC differentiation, endocytosis and actin organization. These peptides or  
 CC proteins are also useful for diagnosing or treating conditions associated  
 CC with inappropriate expression of Bin2, e.g. liver cancer, breast cancer,  
 CC hepatocarcinoma, myeloid and lymphoid leukemias or blood disorders. The  
 CC peptides, nucleic acid sequences or anti-Bin2 antibodies are useful for  
 CC diagnosing inappropriate expression of Bin2 Bin2 is also useful for  
 CC treating disorders associated with excessive Bin1 levels, e.g. liver,  
 CC colorectal, prostate or breast cancers, epithelia cell cancers, melanoma,  
 CC or hyperplastic disease states. The present sequence represents the Bar  
 CC domain of Bin1 used in homology studies with Bin2 protein  
 XX  
 XX SQ Sequence 252 AA;

Query Match 60.7%; Score 683; DB 4; Length 252;  
 Best Local Similarity 61.0%; Pred. No. 7.6e-49;  
 Matches 136; Conservative 32; Mismatches 53; Indels 2; Gaps 2;  
 QY 1 MAE-GRAGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASNYQQAGHGK 59  
 Db 1 MAEMSGKGVYTAGKIASNVQKLTQRAQEKVLQKLGKADETKDEQFCQVQNFNKLTEGTR 60  
 QY 60 LYKDLNFLSAVKVMHSSKEVSETLQETYSSEWDGHEELKAIWNNDLLWEDYEELKAD 119  
 Db 61 LQKDLRTYLASVKAMHEASKLSECLQEVYEPWPGRDEANKIAENNDLLWMDYHQKLV 120  
 QY 120 QAVRTMEIYVAQSEIKERIAKRGKLVYDSDARHLEAVQNA-KKDEAKTAKAEFEFNK 178  
 Db 121 QALLTMDTYLGGFPDIKSIKRIAKRGKLVYDSDARHLYESLQTKKKDEAKIAKEEBLIK 180  
 QY 179 AQTVPEDLNQELLELPILYNSRIGCVYTFONISNLRDVFYR 221  
 Db 181 AOKVFEEMNVDLQELPSLWNSRVGFYVNTFQSIAGLEENFHK 223

RESULT 12  
 ABU89771  
 ID ABU89771 standard; protein; 482 AA.  
 AC ABU89771;  
 DT 10-JUL-2003 (first entry)  
 XX Protein differentially expressed in cardiovascular disease #65.  
 DE Cardiovascular disease; arteriosclerosis; ischaemia; angina pectoris;  
 KW myocardial infarction; cardiast; antiarteriosclerotic; antianginal;  
 KW gene therapy; differential gene expression.  
 XX Homo sapiens.  
 OS WO2003031650-A2.  
 PN 17-APR-2003.  
 PD 02-OCT-2002; 2002WO-EP011034.  
 PF 08-OCT-2001; 2001GB-00024145.  
 PR (FARB ) BAYER AG.  
 PA Munnes M, Gehrman M, Wick M, Schmitz G;  
 PI  
 XX





```

AC AAW47295;
XX
DT 01-JUN-1998 (first entry)
XX
DE DE
DE DE
XX Human Bin1.
XX
KW Human; Box-dependent myc-interacting protein; Bin1; treatment; cancer;
KW hyperplastic disease; neurodegenerative disease; antibody; diagnosis.
XX
OS Homo sapiens.
XX
PN US5723581-A.
XX
PD 03-MAR-1998.
XX
PP 24-MAY-1996; 96US-00652972.
XX
PR 05-MAY-1995; 95US-00435454.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Sakamuro D, Prendergast GC;
XX
DR WPI; 1998-178541/16.
XX N-PSDB; AAV15692.
XX
PT Box-dependent myc-interacting protein Bin1 - useful for, e.g. treating
PT diseases associated with inadequate Bin1 levels such as colon cancer.
XX
XX Claim 1; Col 25-30; 32pp; English.
XX
XX The present sequence is human Box-dependent myc-interacting protein Bin1,
XX which can be used to treat diseases associated with inadequate Bin1
XX levels, e.g. liver, colon or breast cancer or hyperplastic diseases, or
XX diseases where Bin1 activity needs to be inhibited, e.g.
XX neurodegenerative diseases, or to raise antibodies for diagnostic
XX purposes
XX
SQ Sequence 451 AA;
Query Match 59.7%; Score 671.5; DB 2; Length 451;
Best Local Similarity 61.5%; Pred. No. 1.4e-47;
Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;
QY 10 AGLFKQVQKKFRAQEKVLQKLGKAVETKDRFQSQSNFYQQQAEHGKLYKDLKNFLS 69
Db 8 AGKIASNVQKKLTRAQEKVLQKLGKADETKDEQFCQVQNFKNQLTGTRLQKDLRTYLA 67
QY 70 AVKVMHSSKRVSETLOEIIYSSEWDGHEELKAIWVNDLLWEDYEEKLADQAVRTMEIYV 129
Db 68 SVKAMHEASKKLNCECLQEVPEPDPGRDEANKIAENNDLLWMDYHQKLVQALLTMDTYL 127
QY 130 AQPSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAEFEFNKAQTVFEDLNQ 188
Db 128 GQFPDIKRIAKGRKLVYDSARHYESLSQAKKDEAKIAKAEELIKAKQVFEENVV 187
QY 189 ELLEELPILYNSRIGCYVTFIPQNISLRDVFYR 221
Db 188 DLOEELPSLWNSRVGFYNTFQSIAGLEENFHK 220
RESULT 15
AAB36881
ID AAB36881 standard; protein; 451 AA.
XX
AC AAB36881;
XX
DT 22-FEB-2001 (first entry)
XX
DE Human Bin1 protein.
XX
KW Human; Bin1-associated U1-specific; Bau; cancer; gene therapy.
XX

```

```

OS Homo sapiens.
XX
PN US6140465-A.
XX
PD 31-OCT-2000.
XX
PP 29-JUN-1999; 99US-00344889.
XX
PR 29-AUG-1996; 96US-0025482P.
PR 27-AUG-1997; 97US-00919145.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Prendergast GC;
XX
DR WPI; 1998-207038/18.
XX N-PSDB; AAC68498.
XX
PT New isolated Bin1-associated U1-specific protein - which acts as a tumour
PT suppressor, used to develop products for treating e.g. cancers, hyper-
PT plastic disease states or degenerative diseases.
XX
PS Disclosure; Fig 3; 22pp; English.
XX
XX The present invention relates to mammalian Bin1-associated U1-specific
XX (Bau) protein. The Bau protein is useful in the detection, diagnosis and
XX treatment of cancers or other disorders associated with inappropriate
XX Bin1 levels and/or deregulation, deficiency or amplification of the c-Myc
XX oncogenes. The proteins are also useful as antigens for the development
XX of anti-Bau antisera and antibodies to Bau, or to a desired fragment of
XX the Bau protein, as diagnostic reagents, in gene therapy, and in
XX screening and developing chemical compounds or proteins which may be used
XX for the treatment of cancers characterized by Bau or Bin1, which regulate
XX inappropriate MYC levels
XX
SQ Sequence 451 AA;
Query Match 59.7%; Score 671.5; DB 2; Length 451;
Best Local Similarity 61.5%; Pred. No. 1.4e-47;
Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;
QY 10 AGLFKQVQKKFRAQEKVLQKLGKAVETKDRFQSQSNFYQQQAEHGKLYKDLKNFLS 69
Db 8 AGKIASNVQKKLTRAQEKVLQKLGKADETKDEQFCQVQNFKNQLTGTRLQKDLRTYLA 67
QY 70 AVKVMHSSKRVSETLOEIIYSSEWDGHEELKAIWVNDLLWEDYEEKLADQAVRTMEIYV 129
Db 68 SVKAMHEASKKLNCECLQEVPEPDPGRDEANKIAENNDLLWMDYHQKLVQALLTMDTYL 127
QY 130 AQPSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAEFEFNKAQTVFEDLNQ 188
Db 128 GQFPDIKRIAKGRKLVYDSARHYESLSQAKKDEAKIAKAEELIKAKQVFEENVV 187
QY 189 ELLEELPILYNSRIGCYVTFIPQNISLRDVFYR 221
Db 188 DLOEELPSLWNSRVGFYNTFQSIAGLEENFHK 220

```

Search completed: March 4, 2004, 17:26:43  
Job time : 29.4344 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run On: March 4, 2004, 17:46:32 ; Search time 31.9355 Seconds  
(without alignments)  
119.014 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_138\_155

Perfect score: 95

Sequence: 1 RIAGKGRKLVDYDSARHH 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications AA:\*\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	404	9	US-09-879-957-24
2	95	100.0	434	9	US-09-879-957-22
3	95	100.0	451	14	US-10-123-807-4
4	95	100.0	482	15	US-10-116-275-238
5	50	52.6	351	15	US-10-369-493-3387
6	47	49.5	265	15	US-10-369-493-12681
7	45.5	47.9	742	15	US-10-104-047-2962
8	43	45.3	237	15	US-10-369-493-2408
9	43	45.3	284	15	US-10-369-493-2410
10	42	44.2	634	9	US-09-731-221-78
11	42	44.2	906	13	US-10-060-230-19
12	42	44.2	906	13	US-10-060-230-20
13	42	44.2	906	13	US-10-060-230-21
14	42	44.2	906	13	US-10-060-230-22
15	41.5	43.7	593	14	US-10-156-761-14765

16	41	43.2	90	14	US-10-203-081-4	Sequence 4, Appli
17	41	43.2	237	14	US-10-080-170-500	Sequence 500, App
18	41	43.2	657	15	US-10-369-493-12857	Sequence 12857, A
19	41	43.2	771	14	US-10-128-714-3280	Sequence 3280, Ap
20	41	43.2	905	14	US-10-128-714-8280	Sequence 8280, Ap
21	40.5	42.6	511	15	US-10-259-194A-180	Sequence 180, App
22	40	42.1	276	14	US-10-156-761-13731	Sequence 13731, A
23	40	42.1	294	9	US-09-927-738-9	Sequence 9, Appli
24	40	42.1	406	14	US-10-205-219-125	Sequence 125, App
25	40	42.1	414	9	US-09-925-302-706	Sequence 706, App
26	40	42.1	414	9	US-09-927-738-8	Sequence 8, Appli
27	40	42.1	414	15	US-10-264-049-2843	Sequence 2843, Ap
28	40	42.1	436	15	US-10-262-445-6	Sequence 6, Appli
29	40	42.1	440	10	US-09-932-367A-19	Sequence 19, Appl
30	40	42.1	440	15	US-10-369-493-7014	Sequence 7014, Ap
31	39	41.1	30	14	US-10-029-386-27685	Sequence 27685, A
32	39	41.1	315	14	US-10-306-762-187	Sequence 187, App
33	39	41.1	360	9	US-09-847-057-2	Sequence 2, Appli
34	39	41.1	390	15	US-10-369-493-12646	Sequence 12646, A
35	39	41.1	450	10	US-09-949-029-72	Sequence 72, Appl
36	39	41.1	471	14	US-10-128-714-3080	Sequence 3080, Ap
37	39	41.1	565	14	US-10-128-714-8080	Sequence 8080, Ap
38	38.5	40.5	724	14	US-10-211-962-25	Sequence 25, Appl
39	38	40.0	90	15	US-10-630-590-149	Sequence 149, App
40	38	40.0	225	15	US-10-369-493-20139	Sequence 20139, A
41	38	40.0	306	15	US-10-320-797-3084	Sequence 3084, Ap
42	38	40.0	314	15	US-10-369-493-10341	Sequence 10341, A
43	38	40.0	475	9	US-09-925-301-1263	Sequence 1263, Ap
44	38	40.0	632	14	US-10-128-714-8191	Sequence 8191, Ap
45	38	40.0	715	9	US-09-870-759-86	Sequence 86, Appl
46	38	40.0	715	10	US-09-751-708A-86	Sequence 86, Appl
47	38	40.0	757	15	US-10-369-493-16797	Sequence 16797, A
48	38	40.0	950	9	US-09-815-242-11961	Sequence 11961, A
49	38	40.0	1433	14	US-10-224-249-14	Sequence 14, Appl
50	38	40.0	1518	9	US-09-801-368-152	Sequence 152, App
51	38	40.0	1518	15	US-10-369-493-22243	Sequence 22243, A
52	37.5	39.5	636	15	US-10-369-493-17930	Sequence 17930, A
53	37.5	39.5	1249	15	US-10-369-493-6039	Sequence 6039, Ap
54	37	38.9	81	10	US-09-813-153-115	Sequence 115, App
55	37	38.9	89	9	US-09-815-242-11224	Sequence 11224, A
56	37	38.9	89	9	US-09-815-242-11245	Sequence 11245, A
57	37	38.9	110	14	US-10-156-761-7872	Sequence 7872, Ap
58	37	38.9	114	14	US-10-156-761-11982	Sequence 11982, A
59	37	38.9	145	9	US-09-815-242-5551	Sequence 5551, Ap
60	37	38.9	145	9	US-09-815-242-12237	Sequence 12237, A
61	37	38.9	150	15	US-10-369-493-1467	Sequence 1467, Ap
62	37	38.9	338	14	US-10-369-493-21569	Sequence 21569, A
63	37	38.9	342	14	US-10-156-761-9059	Sequence 9059, Ap
64	37	38.9	355	10	US-09-882-227-608	Sequence 608, App
65	37	38.9	381	15	US-10-369-493-10515	Sequence 10515, A
66	37	38.9	474	15	US-10-369-493-19396	Sequence 19396, A
67	37	38.9	513	15	US-10-369-493-13196	Sequence 13196, A
68	37	38.9	577	14	US-10-156-761-12668	Sequence 12668, A
69	37	38.9	642	15	US-10-108-260A-4483	Sequence 4483, Ap
70	37	38.9	840	15	US-10-369-493-3233	Sequence 3233, Ap
71	37	38.9	1139	15	US-10-369-493-5826	Sequence 5826, Ap
72	37	38.9	1277	9	US-09-969-362-3	Sequence 3, Appli
73	36	37.9	89	9	US-09-815-242-10313	Sequence 10313, A
74	36	37.9	89	9	US-09-815-242-11754	Sequence 11754, A
75	36	37.9	89	9	US-09-815-242-13767	Sequence 13767, A
76	36	37.9	130	15	US-10-364-743-83	Sequence 83, Appl
77	36	37.9	170	9	US-09-738-626-5096	Sequence 5096, Ap
78	36	37.9	188	15	US-10-369-493-18256	Sequence 18256, A
79	36	37.9	206	10	US-09-764-891-4189	Sequence 4189, Ap
80	36	37.9	252	15	US-10-369-493-14032	Sequence 14032, A
81	36	37.9	254	14	US-10-441-626-9	Sequence 9, Appli
82	36	37.9	254	14	US-10-441-626-10	Sequence 10, Appl
83	36	37.9	269	15	US-10-264-213-201	Sequence 201, App
84	36	37.9	271	15	US-10-369-493-9095	Sequence 9095, Ap
85	36	37.9	306	15	US-10-369-493-3028	Sequence 3028, Ap
86	36	37.9	317	14	US-10-306-762-230	Sequence 230, App
87	36	37.9	336	15	US-10-369-493-19238	Sequence 19238, A
88	36	37.9	347	14	US-10-156-761-9006	Sequence 9006, Ap

89 Sequence 9212, Ap  
90 Sequence 9453, Ap  
91 Sequence 9458, Ap  
92 Sequence 12394, A  
93 Sequence 2049, Ap  
94 Sequence 413, Appl  
95 Sequence 122, App  
96 Sequence 7, Appl  
97 Sequence 2271, Ap  
98 Sequence 2271, Ap  
99 Sequence 2218, A  
100 Sequence 2218, A  
101 Sequence 10880, A  
102 Sequence 12465, A  
103 Sequence 14716, A  
104 Sequence 15197, A  
105 Sequence 11428, A  
106 Sequence 14391, A  
107 Sequence 22152, A  
108 Sequence 9875, Ap  
109 Sequence 14788, A  
110 Sequence 21698, A  
111 Sequence 11685, A  
112 Sequence 14947, A  
113 Sequence 4, Appl  
114 Sequence 3149, Ap  
115 Sequence 5, Appl  
116 Sequence 11238, A  
117 Sequence 132, App  
118 Sequence 16805, A  
119 Sequence 22621, A  
120 Sequence 12, Appl  
121 Sequence 7, Appl  
122 Sequence 260, App  
123 Sequence 9148, Ap  
124 Sequence 22621, A  
125 Sequence 22622, A  
126 Sequence 7, Appl  
127 Sequence 184, App  
128 Sequence 6, Appl  
129 Sequence 23, Appl  
130 Sequence 14, Appl  
131 Sequence 5006, Ap  
132 Sequence 10948, A  
133 Sequence 1978, Ap  
134 Sequence 5, Appl  
135 Sequence 13344, A  
136 Sequence 54, Appl  
137 Sequence 8173, Ap  
138 Sequence 3747, A  
139 Sequence 3953, Ap  
140 Sequence 338, App  
141 Sequence 137, App  
142 Sequence 13, Appl  
143 Sequence 336, App  
144 Sequence 1519, Ap  
145 Sequence 85, Appl  
146 Sequence 85, Appl  
147 Sequence 84, Appl  
148 Sequence 84, Appl  
149 Sequence 7244, Ap  
150 Sequence 7244, Ap

ALIGNMENTS

RESULT 1  
US-09-879-957-24  
; Sequence 24, Application US/09879957  
; Patent No. US20020034755A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.

Sequence 9212, Ap  
Sequence 9453, Ap  
Sequence 9458, Ap  
Sequence 12394, A  
Sequence 2049, Ap  
Sequence 413, Appl  
Sequence 122, App  
Sequence 7, Appl  
Sequence 2271, Ap  
Sequence 2271, Ap  
Sequence 2218, A  
Sequence 2218, A  
Sequence 10880, A  
Sequence 12465, A  
Sequence 14716, A  
Sequence 15197, A  
Sequence 11428, A  
Sequence 14391, A  
Sequence 22152, A  
Sequence 9875, Ap  
Sequence 14788, A  
Sequence 21698, A  
Sequence 11685, A  
Sequence 14947, A  
Sequence 4, Appl  
Sequence 3149, Ap  
Sequence 5, Appl  
Sequence 11238, A  
Sequence 132, App  
Sequence 16805, A  
Sequence 22621, A  
Sequence 12, Appl  
Sequence 7, Appl  
Sequence 260, App  
Sequence 9148, Ap  
Sequence 22621, A  
Sequence 22622, A  
Sequence 7, Appl  
Sequence 184, App  
Sequence 6, Appl  
Sequence 23, Appl  
Sequence 14, Appl  
Sequence 5006, Ap  
Sequence 10948, A  
Sequence 1978, Ap  
Sequence 5, Appl  
Sequence 13344, A  
Sequence 54, Appl  
Sequence 8173, Ap  
Sequence 3747, A  
Sequence 3953, Ap  
Sequence 338, App  
Sequence 137, App  
Sequence 13, Appl  
Sequence 336, App  
Sequence 1519, Ap  
Sequence 85, Appl  
Sequence 85, Appl  
Sequence 84, Appl  
Sequence 84, Appl  
Sequence 7244, Ap  
Sequence 7244, Ap

HOFFMAN, No. US20020034755A1h  
KAY, Brian K.  
FOWLES, Dana M.  
MCCONNELL, Stephen J.  
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
USING SAME  
NUMBER OF SEQUENCES: 227  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/879,957  
FILING DATE: 13-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,915  
FILING DATE: 03-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 750-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 404 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-09-879-957-24  
Query Match 100.0%; Score 95; DB 9; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RIAGEGRKLVYDSARHH 18  
Db 104 RIAGEGRKLVYDSARHH 121  
RESULT 2  
US-09-879-957-22  
; Sequence 22, Application US/09879957  
; Patent No. US20020034755A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; HOFFMAN, No. US20020034755A1h  
; KAY, Brian K.  
; FOWLES, Dana M.  
; MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; USING SAME  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York

ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60DPCT  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-123-807-4

Query Match 100.0%; Score 95; DB 14; Length 451;  
Best Local Similarity 100.0%; Pred. No. 2.7e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
Db 139 RIAKGRKLVYDSARHH 153

RESULT 4  
US-10-116-275-238  
Sequence 238, Application US/10116275  
Publication No. US20030211476A1  
GENERAL INFORMATION:  
APPLICANT: Elan Pharmaceutical Technology  
APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Brayden, David  
APPLICANT: Byrne, Daragh  
APPLICANT: Lambkin, Imelda  
APPLICANT: Higgins, Lisa  
TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
FILE OF INVENTION: E1067/20087  
CURRENT APPLICATION NUMBER: US/10/116,275  
CURRENT FILING DATE: 2002-10-04  
NUMBER OF SEQ ID NOS: 349  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 238  
LENGTH: 482  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-116-275-238

Query Match 100.0%; Score 95; DB 15; Length 482;  
Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
Db 139 RIAKGRKLVYDSARHH 156

RESULT 5  
US-10-369-493-3387  
Sequence 3387, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28

COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/879,957  
FILING DATE: 13-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,915  
FILING DATE: 03-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELEPHONE: (212) 790-3090  
TELEFAX: (212) 869-8864/9741  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-879-957-22

Query Match 100.0%; Score 95; DB 9; Length 434;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIAKGRKLVYDSARHH 18  
Db 139 RIAKGRKLVYDSARHH 156

RESULT 3  
US-10-123-807-4  
Sequence 4, Application US/10123807  
Publication No. US20030166021A1  
GENERAL INFORMATION:  
APPLICANT: Wistar Institute of Anatomy & Biology  
APPLICANT: Prendergast, George C.  
APPLICANT: Sakamuro, Daitoku  
TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESS: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/123,807  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/445,247  
FILING DATE: 03-Dec-1999  
APPLICATION NUMBER: US 08/870,126  
FILING DATE: 06-JUN-1997

```
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3387
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(351)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3387

Query Match          52.6%; Score 50; DB 15; Length 351;
Best Local Similarity 69.2%; Pred. No. 4.6;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 KRGRKLVYDSAR 16
Db 226 KRGHKLVDYDALR 238

RESULT 6
US-10-369-493-12681
; Sequence 12681, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12681
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(265)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12681

Query Match          49.5%; Score 47; DB 15; Length 265;
Best Local Similarity 60.0%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDSAR 16
Db 151 IKGRNNKLVDYDAMR 165

RESULT 7
US-10-104-047-2962
; Sequence 2962, Application US/10104047
; Publication No. US20030233692A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030233692A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1

; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3387
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2962

Query Match          47.9%; Score 45.5; DB 15; Length 742;
Best Local Similarity 43.5%; Pred. No. 55;
Matches 10; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

QY 1 RIAGRGKLV-----DYDSARHH 18
Db 387 RVGRGRVLLDRAHSDYDSVFHH 409

RESULT 8
US-10-369-493-2408
; Sequence 2408, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2408
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(237)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2408

Query Match          45.3%; Score 43; DB 15; Length 237;
Best Local Similarity 53.3%; Pred. No. 42;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDSAR 16
Db 105 ITRNNHKLVDHDMR 119

RESULT 9
US-10-369-493-2410
; Sequence 2410, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2410
; LENGTH: 284
; TYPE: PRT
```

; ORGANISM: Schizosaccharomyces pombe  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(284)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-2410

Query Match 45.3%; Score 43; DB 15; Length 284;  
Best Local Similarity 53.3%; Pred. No. 51;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IAKGRKLVYDSAR 16  
Db 152 ITRNKHLLDHDAMR 166

RESULT 10  
US-09-731-221-78  
; Sequence 78, Application US/09731221  
; Patent No. US20020018778A1  
; GENERAL INFORMATION:  
; APPLICANT: Caplan, Michael  
; TITLE OF INVENTION: Passive Desensitization  
; FILE REFERENCE: 2002834-0103  
; CURRENT APPLICATION NUMBER: US/09/731,221  
; CURRENT FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 78  
; LENGTH: 634  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Arachis  
US-09-731-221-78

Query Match 44.2%; Score 42; DB 9; Length 634;  
Best Local Similarity 64.3%; Pred. No. 1.7e+02;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 AKGRKLVYDSAR 16  
Db 86 ATRGRQFGDYDDAR 99

RESULT 11  
US-10-060-230-19  
; Sequence 19, Application US/10060230  
; Publication No. US20020173014A1  
; GENERAL INFORMATION:  
; APPLICANT: HILTUNEN, Kalervo  
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA  
; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate  
; FILE REFERENCE: 0365-0528P  
; CURRENT APPLICATION NUMBER: US/10/060,230  
; CURRENT FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 19991667  
; PRIOR FILING DATE: 1999-08-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 906  
; TYPE: PRT  
; ORGANISM: Candida tropicalis  
US-10-060-230-19

Query Match 44.2%; Score 42; DB 13; Length 906;  
Best Local Similarity 46.7%; Pred. No. 2.5e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 IAKGRKLVYDSAR 16

Db 267 VAKGFSEILDYDSR 281

RESULT 12  
US-10-060-230-20  
; Sequence 20, Application US/10060230  
; Publication No. US20020173014A1  
; GENERAL INFORMATION:  
; APPLICANT: HILTUNEN, Kalervo  
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA  
; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate  
; FILE REFERENCE: 0365-0528P  
; CURRENT APPLICATION NUMBER: US/10/060,230  
; CURRENT FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 19991667  
; PRIOR FILING DATE: 1999-08-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 906  
; TYPE: PRT  
; ORGANISM: Candida tropicalis  
US-10-060-230-20

Query Match 44.2%; Score 42; DB 13; Length 906;  
Best Local Similarity 46.7%; Pred. No. 2.5e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 IAKGRKLVYDSAR 16  
Db 267 VAKGFSEILDYDSR 281

RESULT 13  
US-10-060-230-21  
; Sequence 21, Application US/10060230  
; Publication No. US20020173014A1  
; GENERAL INFORMATION:  
; APPLICANT: HILTUNEN, Kalervo  
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA  
; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate  
; FILE REFERENCE: 0365-0528P  
; CURRENT APPLICATION NUMBER: US/10/060,230  
; CURRENT FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 19991667  
; PRIOR FILING DATE: 1999-08-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 906  
; TYPE: PRT  
; ORGANISM: Candida tropicalis  
US-10-060-230-21

Query Match 44.2%; Score 42; DB 13; Length 906;  
Best Local Similarity 46.7%; Pred. No. 2.5e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 IAKGRKLVYDSAR 16  
Db 267 VAKGFSEILDYDSR 281

RESULT 14  
US-10-060-230-22  
; Sequence 22, Application US/10060230  
; Publication No. US20020173014A1  
; GENERAL INFORMATION:  
; APPLICANT: HILTUNEN, Kalervo  
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA

; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate  
; TITLE OF INVENTION: synthesis in genetically modified organisms  
; FILE REFERENCE: 0365-0528P  
; CURRENT APPLICATION NUMBER: US/10/060,230  
; CURRENT FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 19991667  
; PRIOR FILING DATE: 1999-08-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 906  
; TYPE: PRT  
; ORGANISM: Candida tropicalis  
US-10-060-230-22

Query Match 44.2%; Score 42; DB 13; Length 906;  
Best Local Similarity 46.7%; Pred. No. 2.5e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDSAR 16  
Db 267 VAKRFSEILDYDSR 281  
:||||:||||:|

RESULT 15  
US-10-156-761-14765  
; Sequence 14765, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14765  
; LENGTH: 593  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14765

Query Match 43.7%; Score 41.5; DB 14; Length 593;  
Best Local Similarity 72.7%; Pred. No. 1.9e+02;  
Matches 8; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 8 KLVYDSARHH 18  
Db 514 QLIDYDS-RHH 523  
:||||:||||

RESULT 16  
US-10-081-4  
; Sequence 4, Application US/10203081  
; Publication No. US20030082527A1  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Gary K.  
; TITLE OF INVENTION: Viral Cell Death Protein and Uses Therefore  
; FILE REFERENCE: P03909  
; CURRENT APPLICATION NUMBER: US/10/203,081  
; CURRENT FILING DATE: 2002-08-05  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4

; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Snowshoe hare virus  
US-10-203-081-4

Query Match 43.2%; Score 41; DB 14; Length 90;  
Best Local Similarity 37.5%; Pred. No. 32;  
Matches 6; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 RIAKRGKLVYDSAR 16  
Db 47 RVSGRQIILUESGR 62  
:||||:||||:|

RESULT 17  
US-10-080-170-500  
; Sequence 500, Application US/10080170  
; Publication No. US20030129601A1  
; GENERAL INFORMATION:  
; APPLICANT: COLE, S.T.  
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
; FILE REFERENCE: 03495.0218  
; CURRENT APPLICATION NUMBER: US/10/080,170  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: 60/270,123  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 652  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 500  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-10-080-170-500

Query Match 43.2%; Score 41; DB 14; Length 237;  
Best Local Similarity 52.9%; Pred. No. 89;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 IAKRGKLVYDSARHH 18  
Db 79 LTKRGKALDDILAARDH 95  
:||||:||||

RESULT 18  
US-10-369-493-12857  
; Sequence 12857, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 12857  
; LENGTH: 657  
; TYPE: PRT  
; ORGANISM: Aspergillus nidulans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(657)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-12857

```

; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 6603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8280
; LENGTH: 905
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8280

Query Match      43.2%; Score 41
Best Local Similarity 50.0%; Pred. No. 1
Matches      8; Conservative      3; Mismatch      4

QY      3 AKGRKLVVDYSARHH 18
       :|||:::|||:
Db      362 SSRGAGLTVDVSARHN 377

RESULT 21
US-10-259-194A-180
; Sequence 180, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghahremanian, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyuki
; APPLICANT: Krepe, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Darroll
; APPLICANT: Ricke, Nathrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND C
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,19
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: PatentList.pl version 3.0.4 (C
; SEQ ID NO 180
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Cryza eativa
US-10-259-194A-180

Query Match      42.6%; Score 40.
Best Local Similarity 53.3%; Pred. No. 1
Matches      8; Conservative      4; Mismatch      4

QY      4 KRGKRLVVDYSARHH 18
       :|||:::|||:
Ddb     141 RRGAEIDF-SARHH 154

RESULT 22
US-10-156-761-13731
; Sequence 13731, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, TARUHO

```



APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 13731  
LENGTH: 276  
TYPE: PRT  
ORGANISM: Streptomyces avermitilis  
US-10-156-761-13731

Query Match 42.1%; Score 40; DB 14; Length 276;  
Best Local Similarity 53.6%; Pred. No. 1.5e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RKLVYDSARH 17  
DB 83 RELADVDAARH 93

## RESULT 23

US-09-927-738-9  
Sequence 9, Application US/09927738  
Patent No. US20020076799A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongwen  
TITLE OF INVENTION: Compositions and Methods of modulating TGF-B signaling  
FILE REFERENCE: 17633/1082  
CURRENT APPLICATION NUMBER: US/09/927,738  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: US 60/119786  
PRIOR FILING DATE: 1999-02-11  
PRIOR APPLICATION NUMBER: PCT/US00/03561  
PRIOR FILING DATE: 2000-02-11  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Clone S1+12 -5 protein  
US-09-927-738-9

Query Match 42.1%; Score 40; DB 9; Length 294;  
Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KGRKLVYDSDA 15  
DB 200 RRSRLVDYENA 211

## RESULT 24

US-10-205-219-125  
Sequence 125, Application US/10205219  
Publication No. US20030138603A1  
GENERAL INFORMATION:  
APPLICANT: Warner-Lambert Company  
APPLICANT: Lee, Kevin  
APPLICANT: Dixon, Alistair  
APPLICANT: Brooksbank, Robert  
APPLICANT: Pincock, Robert  
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain

FILE REFERENCE: WL-A-018200  
CURRENT APPLICATION NUMBER: US/10/205,219  
CURRENT FILING DATE: 2002-07-24  
PRIOR APPLICATION NUMBER: GB 0118354.0  
PRIOR FILING DATE: 2001-07-27  
NUMBER OF SEQ ID NOS: 197  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 125  
LENGTH: 406  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: SNX6  
US-10-205-219-125

Query Match 42.1%; Score 40; DB 14; Length 406;  
Best Local Similarity 58.3%; Pred. No. 2.3e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KGRKLVYDSDA 15  
DB 312 RRSRLVDYENA 323

## RESULT 25

US-09-925-302-706  
Sequence 706, Application US/09925302  
Patent No. US2002004941A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA104  
CURRENT APPLICATION NUMBER: US/09/925,302  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05918  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 896  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 706  
LENGTH: 414  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-302-706

Query Match 42.1%; Score 40; DB 9; Length 414;  
Best Local Similarity 58.3%; Pred. No. 2.3e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KGRKLVYDSDA 15  
DB 320 RRSRLVDYENA 331

## RESULT 26

US-09-927-738-8  
Sequence 8, Application US/09927738  
Patent No. US20020076799A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongwen  
TITLE OF INVENTION: Compositions and Methods of modulating TGF-B signaling  
FILE REFERENCE: 17633/1082  
CURRENT APPLICATION NUMBER: US/09/927,738  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: US 60/119786  
PRIOR FILING DATE: 1999-02-11  
PRIOR APPLICATION NUMBER: PCT/US00/03561  
PRIOR FILING DATE: 2000-02-11  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 414

```
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Clone S1-12 -2 protein Sequence
US-09-927-738-8

Query Match          42.1%; Score 40; DB 9; Length 414;
Best Local Similarity 58.3%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 KGRKRLVDYDSA 15
DB      320 RRSRLVDYENA 331

RESULT 27
US-10-264-049-2843
; Sequence 2843, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2843
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2843

Query Match          42.1%; Score 40; DB 15; Length 414;
Best Local Similarity 58.3%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 KGRKRLVDYDSA 15
DB      320 RRSRLVDYENA 331

RESULT 28
US-10-262-445-6
; Sequence 6, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Shionit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Giot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zhong, Haibong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
```

```
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 6
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-445-6

Query Match          42.1%; Score 40; DB 15; Length 436;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 KGRKRLVDYDSA 15
DB      342 RRSRLVDYENA 353

RESULT 29
US-09-932-367A-19
; Sequence 19, Application US/09932367A
; Publication No. US20030027152A1
; GENERAL INFORMATION:
; APPLICANT: RHODES, Simon J.
; APPLICANT: BRIDWELL, Jeanne L.
; APPLICANT: MEIER, Bradley C.
; APPLICANT: PARKER, Gretchen E.
; APPLICANT: PRICE, Jeffrey R.
; APPLICANT: SHOWALTER, Aaron D.
; APPLICANT: SLOOP, Kyle W.
; TITLE OF INVENTION: GENERATION OF DIAGNOSTIC TOOLS TO ASSAY THE HUMAN
; FILE REFERENCE: LHX3/P-LIM/LIN-3 FACTOR
; CURRENT APPLICATION NUMBER: US/09/932,367A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/04424
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/121,110
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-932-367A-19

Query Match          42.1%; Score 40; DB 10; Length 440;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
```

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 VDYDSARHH 18  
| | | | |  
Db 269 VDYDSFSHH 277

## RESULT 30

US-10-369-493-7014  
; Sequence 7014, Application US/10369493  
; Publication No. US20030233675A1

## GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 7014  
; LENGTH: 830  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-369-493-7014

Query Match 42.1%; Score 40; DB 15; Length 830;

Best Local Similarity 46.2%; Pred. No. 4.9e+02;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 KKGRLVDYDSAR 16  
| | | | |  
Db 88 KKGHEFVDFDAVR 100

## RESULT 31

US-10-029-386-27685  
; Sequence 27685, Application US/10029386  
; Publication No. US20030194704A1

## GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 27685  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL137071.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4  
US-10-029-386-27685

Query Match 41.1%; Score 39; DB 14; Length 30;

Best Local Similarity 50.0%; Pred. No. 21;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IAKGRKLVYDSARH 17  
| | | | |  
Db 15 ISKEDRELFLSSQRH 30

## RESULT 32

US-10-306-762-187  
; Sequence 187, Application US/10306762  
; Publication No. US20030187220A1

## GENERAL INFORMATION:

; APPLICANT: Park, Frances  
; APPLICANT: Gajiwala, Ketan S.  
; APPLICANT: Buchanan, Sean Grant  
; APPLICANT: Sauder, J. Michael  
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN  
; TITLE OF INVENTION: MONONUCLEOTIDE BINDING PROTEIN (FMNBP)  
; FILE REFERENCE: 52498-20011.00  
; CURRENT APPLICATION NUMBER: US/10/306,762  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: US 60/334,132  
; PRIOR FILING DATE: 2001-11-28  
; NUMBER OF SEQ ID NOS: 242  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 187  
; LENGTH: 315  
; TYPE: PRT  
; ORGANISM: D. melanogaster (18859731)  
US-10-306-762-187

Query Match 41.1%; Score 39; DB 14; Length 315;

Best Local Similarity 52.9%; Pred. No. 2.5e+02;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 IAKGRKLVYDSARH 18  
| | | | |  
Db 257 IEKYLCLVDYDNAPHN 273

## RESULT 33

US-09-847-057-2  
; Sequence 2, Application US/09847057  
; Patent No. US20020004943A1

## GENERAL INFORMATION:

; APPLICANT: AGRINOMICS, LLC.  
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A PAGODA PHENOTYPE IN PLANT  
; FILE REFERENCE: PAGODA  
; CURRENT APPLICATION NUMBER: US/09/847,057  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-847-057-2

Query Match 41.1%; Score 39; DB 9; Length 360;

Best Local Similarity 43.8%; Pred. No. 2.9e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RIAGRKLVYDSAR 16  
| | | | |  
Db 339 RLTKKSAYLVDFESCR 354

## RESULT 34

US-10-369-493-12646  
; Sequence 12646, Application US/10369493  
; Publication No. US20030233675A1

## GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 12646  
 ; LENGTH: 390  
 ; TYPE: PRT  
 ; ORGANISM: Aspergillus nidulans  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)...(390)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-369-493-12646

Query Match 41.1%; Score 39; DB 15; Length 390;  
 Best Local Similarity 51.8%; Pred. No. 3.2e+02;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GRKLVYDSARHH 18  
 Db 16 GRILADFTTRRH 28

## RESULT 35

US-09-949-029-72  
 ; Sequence 72, Application US/09949029  
 ; Publication No. US20030134278A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Karpen, G.H.  
 ; APPLICANT: Dobie, K.W.  
 ; APPLICANT: Kennedy, C.D.  
 ; APPLICANT: Velasco, V.M.  
 ; APPLICANT: McGrath, T.L.  
 ; APPLICANT: Weko, J.  
 ; APPLICANT: Patterson, R.W.

; TITLE OF INVENTION: Identification of chromosome inheritance modifiers in Drosophila  
 ; FILE REFERENCE: 1211.015US1  
 ; CURRENT APPLICATION NUMBER: US/09/949,029  
 ; CURRENT FILING DATE: 2001-09-07  
 ; PRIOR APPLICATION NUMBER: US 60/231,178  
 ; PRIOR FILING DATE: 2000-09-07  
 ; NUMBER OF SEQ ID NOS: 149

; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 72  
 ; LENGTH: 450  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-09-949-029-72

Query Match 41.1%; Score 39; DB 10; Length 450;  
 Best Local Similarity 38.9%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RIAGRGRKLVYDSARHH 18  
 Db 217 KLARYLRLRWDARHH 234

## RESULT 36

US-10-128-714-3080  
 ; Sequence 3080, Application US/10128714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Hu, Wengqi  
 ; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Eroshkin, Alexey M  
 ; APPLICANT: Lemieux, Sebastien M  
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 ; TITLE OF INVENTION: Methods of Use  
 ; FILE REFERENCE: 10182-018-999  
 ; CURRENT APPLICATION NUMBER: US/10/128,714  
 ; CURRENT FILING DATE: 2002-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/285,697  
 ; PRIOR FILING DATE: 2001-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/287,066  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3080  
 ; LENGTH: 471  
 ; TYPE: PRT  
 ; ORGANISM: Aspergillus fumigatus  
 US-10-128-714-3080

Query Match 41.1%; Score 39; DB 14; Length 471;  
 Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 RGRKLVYDSARHH 18  
 Db 425 RGRKRFEDKDSKEH 438

## RESULT 37

US-10-128-714-8080  
 ; Sequence 8080, Application US/10128714  
 ; Publication No. US20030119013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Hu, Wengqi  
 ; APPLICANT: Tishkoff, Daniel  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Eroshkin, Alexey M  
 ; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 ; TITLE OF INVENTION: Methods of Use  
 ; FILE REFERENCE: 10182-018-999  
 ; CURRENT APPLICATION NUMBER: US/10/128,714  
 ; CURRENT FILING DATE: 2002-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/285,697  
 ; PRIOR FILING DATE: 2001-04-23  
 ; PRIOR APPLICATION NUMBER: US 60/287,066  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/295,890  
 ; PRIOR FILING DATE: 2001-06-05  
 ; PRIOR APPLICATION NUMBER: US 60/303,899  
 ; PRIOR FILING DATE: 2001-07-09  
 ; PRIOR APPLICATION NUMBER: US 60/316,362  
 ; PRIOR FILING DATE: 2001-08-31  
 ; NUMBER OF SEQ ID NOS: 8603  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 8080  
 ; LENGTH: 565  
 ; TYPE: PRT  
 ; ORGANISM: Aspergillus fumigatus  
 US-10-128-714-8080

Query Match 41.1%; Score 39; DB 14; Length 565;  
 Best Local Similarity 50.0%; Pred. No. 4.7e+02;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 RGRKLVYDSARHH 18  
| | | | | | | | | |  
Db 519 RGRKFEDKSKRHH 532

RESULT 38  
US-10-211-962-25  
; Sequence 25, Application US/10211962  
; Publication No. US20030082640A1  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/10/211,962  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US/09/562,737  
; PRIOR FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 724  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-10-211-962-25

Query Match 40.5%; Score 38.5; DB 14; Length 724;  
Best Local Similarity 56.2%; Pred. No. 7.4e+02;  
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Qy 4 KRG---RKLVDYDSAR 16  
| | | | | | | | | |  
Db 428 KRGFVKRLFDYDTK 443

RESULT 39  
US-10-630-590-149  
; Sequence 149, Application US/10630590  
; Publication No. US20040018487A1  
; GENERAL INFORMATION:  
; APPLICANT: Lu, Peter  
; APPLICANT: Schweizer, Johannes  
; APPLICANT: Diaz-Sarmiento, Chamorro Samoa  
; APPLICANT: Belmares, Michael P.  
; TITLE OF INVENTION: METHODS OF DIAGNOSING CERVICAL CANCER  
; FILE REFERENCE: VITA-008  
; CURRENT APPLICATION NUMBER: US/10/630,590  
; CURRENT FILING DATE: 2003-07-29  
; PRIOR APPLICATION NUMBER: 60/409,298  
; PRIOR FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/450,464  
; PRIOR FILING DATE: 2003-02-27  
; PRIOR APPLICATION NUMBER: US 02/24655  
; PRIOR FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: 60/309,841  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: 60/360,061  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 10/080,273  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: 60/269,523  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: 09/710,059  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 149  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-630-590-149

Query Match 40.0%; Score 38; DB 15; Length 90;  
Best Local Similarity 62.5%; Pred. No. 98;  
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IAKRGRKLVYDSAR 15  
| | | | | | | | | |  
Db 55 LAVNGRPLVDSLSDSA 70

RESULT 40  
US-10-369-493-20139  
; Sequence 20139, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52082)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 20139  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: No. US20030233675A1toc punctiforme  
US-10-369-493-20139

Query Match 40.0%; Score 38; DB 15; Length 225;  
Best Local Similarity 56.2%; Pred. No. 2.6e+02;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 AKRGRKLVYDSARHH 18  
| | | | | | | | | |  
Db 98 AKYGDKLWDTIAKPH 113

RESULT 41  
US-10-320-797-3084  
; Sequence 3084, Application US/10320797  
; Publication No. US20040014955A1  
; GENERAL INFORMATION:  
; APPLICANT: Eroshkin, Alexey M.  
; APPLICANT: Zamudio, Carlos  
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND  
; TITLE OF INVENTION: METHODS OF USE  
; FILE REFERENCE: 10182-021-999  
; CURRENT APPLICATION NUMBER: US/10/320,797  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: 60/341,261  
; PRIOR FILING DATE: 2001-12-17  
; NUMBER OF SEQ ID NOS: 3361  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3084  
; LENGTH: 306  
; TYPE: PRT  
; ORGANISM: Cryptococcus neoformans  
US-10-320-797-3084

Query Match 40.0%; Score 38; DB 15; Length 306;  
Best Local Similarity 50.0%; Pred. No. 3.6e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 RGRKLVYDSARHH 18  
| | | | | | | | | |  
Db 8 RDIQFVDYDNRVRFH 21

## RESULT 42

US-10-369-493-10341  
; Sequence 10341, Application US/10369493  
; Publication No. US20030233675A1

## GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

FILE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 10341

LENGTH: 314

TYPE: PRT

ORGANISM: *Cytophaga hutchinsonii*

US-10-369-493-10341

Query Match 40.0%; Score 38; DB 15; Length 314;

Best Local Similarity 66.7%; Pred. No. 3.7e+02;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 RKLVDYDSARHH 18

Db 303 RDLVDYTIARKH 314

## RESULT 43

US-09-925-301-1263

; Sequence 1263, Application US/09925301

; Patent No. US20020052308A1

## GENERAL INFORMATION:

APPLICANT: Rosen et al.

FILE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA106

CURRENT APPLICATION NUMBER: US/09/925,301

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05882

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1694

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 1263

LENGTH: 475

TYPE: PRT

ORGANISM: *Homo sapiens*

FEATURE:

NAME/KEY: SITE

LOCATION: (249)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-301-1263

Query Match 40.0%; Score 38; DB 9; Length 475;

Best Local Similarity 43.8%; Pred. No. 5.7e+02;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RIARGRKLVYDSAR 16

Db 276 KIFQGESPVYDGR 291

## RESULT 44

US-10-128-714-8191

; Sequence 8191, Application US/10128714

Publication No. US20030119013A1

## GENERAL INFORMATION:

APPLICANT: Jiang, Bo

APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos

APPLICANT: Eroshkin, Alexey M

APPLICANT: Lemieux, Sebastien M

FILE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and

METHODS OF USE

FILE REFERENCE: 10182-018-999

CURRENT APPLICATION NUMBER: US/10/128,714

CURRENT FILING DATE: 2002-04-23

PRIOR APPLICATION NUMBER: US 60/285,697

PRIOR FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: US 60/287,066

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 60/295,890

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: US 60/303,899

PRIOR FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/316,362

PRIOR FILING DATE: 2001-08-31

NUMBER OF SEQ ID NOS: 8603

SOFTWARE: Patent in version 3.1

SEQ ID NO 8191

LENGTH: 632

TYPE: PRT

ORGANISM: *Aspergillus fumigatus*

US-10-128-714-8191

Query Match 40.0%; Score 38; DB 14; Length 632;

Best Local Similarity 61.5%; Pred. No. 7.8e+02;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AKRGKRLVYDSA 15

Db 5 AKPRKLLDDSS 17

## RESULT 45

US-09-870-759-86

; Sequence 86, Application US/09870759

; Patent No. US20020177551A1

## GENERAL INFORMATION:

APPLICANT: Terman, David S

FILE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

FILE REFERENCE: 870759

CURRENT APPLICATION NUMBER: US/09/870,759

CURRENT FILING DATE: 2002-01-14

PRIOR APPLICATION NUMBER: US 60/208,128

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 166

SOFTWARE: Patent in version 3.1

SEQ ID NO 86

LENGTH: 715

TYPE: PRT

ORGANISM: *Staphylococcus aureus*

US-09-870-759-86

Query Match 40.0%; Score 38; DB 9; Length 715;

Best Local Similarity 50.0%; Pred. No. 8.8e+02;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 LVDYDSARHH 18

Db 264 ITSVDSSKHH 273

## RESULT 46

US-09-751-708A-86

; Sequence 86, Application US/09751708A

; Publication No. US20030157113A1

```
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-751-708A-86

Query Match          40.0%; Score 38; DB 10; Length 715;
Best Local Similarity 50.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 LVDYDSARHH 18
   : |||:|
Db 264 ITSVDSSKHH 273

RESULT 47
US-10-369-493-16797
; Sequence 16797, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16797
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-369-493-16797

Query Match          40.0%; Score 38; DB 15; Length 757;
Best Local Similarity 50.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 IAKGRKLVYD 13
   : |||:|:|
Db 364 LEKGRRLIGWD 375

RESULT 48
US-09-815-242-11961
; Sequence 11961, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
```

```
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11961
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11961

Query Match          40.0%; Score 38; DB 9; Length 950;
Best Local Similarity 46.2%; Pred. No. 1.2e+03;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 5 RGRKLVYDSARH 17
   : |||:|:|
Db 165 RGRKLVNWDTKLH 177

RESULT 49
US-10-224-249-14
; Sequence 14, Application US/10224249
; Publication No. US20030087867A1
; GENERAL INFORMATION:
; APPLICANT: Vogels, Ronald V.
; APPLICANT: Verlinden, Stefan F.F.
; TITLE OF INVENTION: Gene therapy for enhancing and/or inducing angiogenesis
; FILE REFERENCE: 2183-5233US
; CURRENT APPLICATION NUMBER: US/10/224,249
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: PCT/NL00/00482
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: EP 99202263.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/143,101
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: CHAIN
; LOCATION: (1)...(1433)
; OTHER INFORMATION: Human nitric oxide synthase
US-10-224-249-14

Query Match          40.0%; Score 38; DB 14; Length 1433;
Best Local Similarity 52.5%; Pred. No. 1.8e+03;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 2 IAKGRKLVYD--YDSA 15
   : |||:|:|
Db 65 LAVNGRELVDLSYDSA 80

RESULT 50
```

US-09-801-368-152  
; Sequence 152, Application US/09801368  
; Patent No. US20020128250A1  
; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert  
; APPLICANT: Cali, Brian  
; APPLICANT: Hecht, Peter  
; APPLICANT: Holtzman, Doug  
; APPLICANT: Madden, Kevin  
; APPLICANT: Maxon, Mary  
; APPLICANT: Milne, Todd  
; APPLICANT: No. US20020128250A1man, Thea  
; APPLICANT: Royer, John  
; APPLICANT: Salama, Sofie  
; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/801,368  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 152  
; LENGTH: 1518  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-152

Query Match 40.0%; Score 38; DB 9; Length 1518;  
Best Local Similarity 70.0%; Pred. No. 2e+03;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 KRGRKLVVDYD 13  
| | | | |  
Db 1451 KFGKRVVEYD 1460

Search completed: March 4, 2004, 17:55:17  
Job time : 33.9355 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:24:55 ; Search time 23.7096 Seconds  
(without alignments)  
1228.072 Million cell updates/sec

Title: US-10-069-540A-2

Perfect score: 2855

Sequence: 1 MAEGKAGGAGLFAKQVOKK.....NNLTAPQPEVSTENPQL 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	816.5	28.6	434	US-08-630-915A-22	Sequence 22, Appl
2	793.5	27.8	451	US-08-652-972A-4	Sequence 4, Appl
3	793.5	27.8	451	US-08-870-126-4	Sequence 4, Appl
4	793.5	27.8	451	US-09-445-247-4	Sequence 4, Appl
5	792	27.7	451	US-08-435-454-4	Sequence 6, Appl
6	792	27.7	451	US-08-919-145-6	Sequence 6, Appl
7	792	27.7	451	US-09-344-883-6	Sequence 6, Appl
8	792	27.7	451	PCT-US96-06231A-4	Sequence 4, Appl
9	675	23.6	404	US-08-630-915A-24	Sequence 24, Appl
10	243	8.5	904	US-09-976-594-615	Sequence 615, App
11	207.5	7.3	1324	US-09-645-456A-13	Sequence 13, Appl
12	207.5	7.3	1324	US-09-425-324A-13	Sequence 13, Appl
13	207.5	7.3	1324	US-09-645-791-13	Sequence 13, Appl
14	205	7.2	1353	US-09-645-456A-11	Sequence 11, Appl
15	205	7.2	1353	US-09-425-324A-11	Sequence 11, Appl
16	205	7.2	1353	US-09-645-791-11	Sequence 11, Appl
17	201.5	7.1	1298	US-09-645-456A-14	Sequence 14, Appl
18	201.5	7.1	1298	US-09-425-324A-14	Sequence 14, Appl
19	201.5	7.1	1298	US-09-645-791-14	Sequence 14, Appl
20	199.5	7.0	1237	US-09-688-188B-14	Sequence 14, Appl
21	199.5	7.0	1237	US-09-291-417D-14	Sequence 14, Appl
22	199.5	7.0	1332	US-09-645-456A-9	Sequence 9, Appl
23	199.5	7.0	1332	US-09-425-324A-9	Sequence 9, Appl
24	199.5	7.0	1332	US-09-645-791-9	Sequence 9, Appl
25	198.5	7.0	1231	US-08-714-741-41	Sequence 41, Appl
26	197	6.9	1360	US-09-393-569-2	Sequence 2, Appl
27	197	6.9	1360	US-09-579-664B-14	Sequence 14, Appl

Sequence 34, Appl  
Sequence 34, Appl  
Sequence 34, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 8, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 5095, Ap

## ALIGNMENTS

### RESULT 1

US-08-630-915A-22

Sequence 22, Application US/08630915A

Patent No. 5409820

GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.

APPLICANT: HOFFMAN, No. 6309820h

APPLICANT: KAY, Brian K.

APPLICANT: FOWLES, Dana M.

APPLICANT: MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL

TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND

TITLE OF INVENTION: USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,915A

FILING DATE: 03-APR-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Misrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 434 amino acids

TYPE: amino acid

STANDARDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-630-915A-22

Query Match 28.6%; Score 816.5; DB 4; Length 434;  
Best Local Similarity 45.1%; Pred. No. 3.3e-47;  
Matches 193; Conservative 47; Mismatches 105; Indels 83; Gaps 11;

QY 1 MAE-GRAGGAAGLFAKQVOKKSRQAQKVLQKLGKAVETKDRPQASNFYQQAGCHK 59  
Db 1 MAEMGSGVYTAGIASNQVOKLTRAQKVLQKLGKADETKDEQFQCVQNFKNQLTETR 60  
QY 60 LYKDLNFKLSAVKMHSSKRVSETLQETIYSSEWDGHEELKAIWVNDLLWEDYEELAD 119  
Db 61 LOKDLRTYLASVAMHSAKSLSECLQEVYEPENFGROEANKIAENNDLLWMDYHQKLV 120  
QY 120 QAVRTMEIYVAQSEIKERIAKGRKLVYDSARHLEAVQNA-KQBAKTAKAEEEPNK 178  
Db 121 QALLTMDYTLQGPFDIKSRIAKGRKLVYDSARHLEAVQNA-KQBAKTAKAEEELIK 180  
QY 179 AQTVFEDLNLEELPILVNSRIGCVYVTFQNIENLSDVYREMSKLNHNLVYVMSKLE 238  
Db 181 AQVFEEMVNDLQBELPILVNSRIGCVYVTFQNIENLSDVYREMSKLNHNLVYVMSKLE 240  
QY 239 QKHSNKVYVYKGLSSSRSLVSPVTRATVSSPLTSPTSLKSESESVSATEDL 298  
Db 241 QKHSNTFTVKAQPSD-----NAPEKGNKSPSP----- 268  
QY 299 APDAAQEDNSDKELLEEEIEKEGSEASSSE---EDDPLACNGPAQAQPSPTTERAK 355  
Db 269 PPD-----GSPATPEIRVNHPEFA-----SGASFGATIPKSP 302  
QY 356 SQBEVLSPSTSPFGGALSPGQSSSATE-----VVLRTTRAS-----EGSEQPKKAS 405  
Db 303 SQ-----PAEASEVVGGAQEPGETAASEATSSSLPAVVVETFSATVNGAVEGS-----AG 352  
QY 406 IQRTSAPP 413  
Db 353 TGRDLPLP 360

## RESULT 2

US-08-972A-4  
; Sequence 4, Application US/08652972A  
; Patent No. 5723581  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; APPLICANT: Sakamuro, Daitoku  
; TITLE OF INVENTION: Murine and Human Box-Dependent  
; TITLE OF INVENTION: MYC-Interacting Protein (Bin1) and Uses Therefor  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/652,972A  
; FILING DATE: 24-MAY-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,454  
; FILING DATE: 05-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: WST60BUSA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-972A-4

Query Match 27.8%; Score 793.5; DB 1; Length 451;  
Best Local Similarity 36.6%; Pred. No. 1.2e-45;  
Matches 193; Conservative 66; Mismatches 152; Indels 117; Gaps 11;  
QY 10 AGLFAKQVOKKSRQAQKVLQKLGKAVETKDRPQASNFYQQAGCHKLYKDLNFKLS 69  
Db 8 AKIASNVQKLTTRAQKVLQKLGKADETKDEQFQCVQNFKNQLTETRQKDLRTYLA 67  
QY 70 AVKWHSSKRVSETLQETIYSSEWDGHEELKAIWVNDLLWEDYEELADQAVRTMEIYV 129  
Db 68 SYKAMHSAKSLSECLQEVYEPENFGROEANKIAENNDLLWMDYHQKLVQALLTMDYTL 127  
QY 130 AQSIEIKERIAKGRKLVYDSARHLEAVQNA-KQBAKTAKAEEEFNKAAQTVFEDLNQ 188  
Db 128 GQFPDIKSRIAKGRKLVYDSARHLEAVQNA-KQBAKTAKAEEELIKAAQVFEEMNV 187  
QY 189 ELLEELPILVNSRIGCVYVTFQNIENLSDVYREMSKLNHNLVYVMSKLEKQHSNKVYV 248  
Db 188 DQBELPILVNSRIGCVYVTFQNIENLSDVYREMSKLNHNLVYVMSKLEKQHSNKVYV 247  
QY 249 KGLSSSRSLVSPVTRATVSSPLTSPTSLKSESESVSATEDLAPDAAQEDN 308  
Db 248 K-----AOPRKSILFSLRRKNSDNAPAKGNKSP----- 278  
QY 309 SEIKELLEEEIEKEGSEASSSEEDDPLACNGPAQAQPSPTTERAKSQBEVLSPSTPS 368  
Db 279 -----SPDGSPEATPEIRVNH-----PE 298  
QY 369 PGGALSPGQSSSATEVVLRTTRASGEGSEQFKRASIQRTSAPPSPRPPPRATASPRPS 428  
Db 299 PAGGATPGA-----TLPKSPQFAEASEVAGGTQPAAGAQEPGETAASEAA 344  
QY 429 SGNIPS-----SPTASGGSPSPRASLTGT-----TASPTSLVSPN 467  
Db 345 SSSLPAVVVETFPATVNGVEGSG--AGRLDLPFGFMFKVQAQHDYATDTDLQKAG 402  
QY 468 PEPPEKPVTRTPEAKE-----NENIHQNPE-ELCTS--PTLMTSQV 505  
Db 403 DIVLVTFQNPBEQDEGLMGVYKESDWNQHKLEKCRGVFFENFTEV 450

## RESULT 3

US-08-870-126-4  
; Sequence 4, Application US/08870126  
; Patent No. 6048702  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; APPLICANT: Sakamuro, Daitoku  
; TITLE OF INVENTION: Murine and Human Box-Dependent  
; TITLE OF INVENTION: MYC-Interacting Protein (Bin1) and Uses Therefor  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/870,126  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/435,454
; FILING DATE: 05-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,972
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: MST60CUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-870-126-4

Query Match 27.8%; Score 793.5; DB 3; Length 451;
Best Local Similarity 36.6%; Pred. No. 1.2e-45;
Matches 193; Conservative 66; Mismatches 152; Indels 117; Gaps 11;

QY 10 AGLFAQVQKFSRAQKVLQKLGKAVETKDERFEQSASNFYQQAAGHKLKYLKLNFLS 69
DB 8 AGKIASNVQKLTQAEKVLQKLGKADETKDEQECVQNFKNQLTETGLQKDLRTYLA 67
QY 70 AVKWHSESKRVSETLQETYSSEWDGHEELKAIYVNNDLWEDYEELKADQAVRTMEIYV 129
DB 68 SVKAMEHASKKLNCEQLQEVPEDPWGRDEANKIAENNDLLWMDYHOKLVQALLTMDTYL 127
QY 130 AQSFIKRIAKRGKLVYDYSARHHLKAVQNA-KDEAKTAKAEEFNKAQTVFEDLNQ 188
DB 128 GQFPDIKSRIAKRGKLVYDYSARHHLKAVQNA-KDEAKTAKAEEFNKAQTVFEDLNQ 187
QY 189 ELLEELPILYNSRIGCVYTFQNIENLSDVFRMSKLNHLYEVMSKLEKQHSNKVYV 248
DB 188 DLQELPILYNSRIGCVYTFQNIENLSDVFRMSKLNHLYEVMSKLEKQHSNKVYV 247
QY 249 KGLSSSRSLVISPPVTRATVSSPLTSPTSLKSESESVSATDLAPDAAGEDN 308
DB 248 K-----AQPRTKSLFSLRRKNSDNAPAKGNKSP-----278
QY 309 SEIKELLEEEIEKEGSEASSEEDDPLPACNGPAQAQPSPTTTERAKSQEVLPSSTPS 368
DB 279 -----SPDGSFAATPEIRVNHE-----PE 298
QY 369 PGGALSPGQPSSTATEVVLRTTASGESEQPKKASIQRTSAPSRPPPPRATASPPS 428
DB 299 PAGGATPGA-----TLPKSPQAEASEVAGGTQPAAGAQEPGETAASEAA 344
QY 429 SGNIPS-----SPTASGGSGPTSPRASLTG-----TASPRTSLEVSPN 467
DB 345 SSSLPAVVVETPPATVNGTVEGSG--AGRLDLPFGFMFKVQAQHDYATATDTDELQKAG 402
QY 468 PEPPEKPVRTPEAKE-----NENIHQNP-E-ELCTS--PTLMTSOV 505
DB 403 DVVLVFPQNPEDQEGWLMGVKESDWNQHKLEKRCGVFPENFTERV 450

```

```

RESULT 4
US-09-445-247-4
; Sequence 4, Application US/09445247
; Patent No. 6410238
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy & Biology
; Prendergast, George C.
; Sakamuro, Daitoku
; TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESS: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/445,247
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/870,126
; FILING DATE: 06-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST60DPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
; US-09-445-247-4

Query Match 27.8%; Score 793.5; DB 4; Length 451;
Best Local Similarity 36.6%; Pred. No. 1.2e-45;
Matches 193; Conservative 66; Mismatches 152; Indels 117; Gaps 11;

QY 10 AGLFAQVQKFSRAQKVLQKLGKAVETKDERFEQSASNFYQQAAGHKLKYLKLNFLS 69
DB 8 AGKIASNVQKLTQAEKVLQKLGKADETKDEQECVQNFKNQLTETGLQKDLRTYLA 67
QY 70 AVKWHSESKRVSETLQETYSSEWDGHEELKAIYVNNDLWEDYEELKADQAVRTMEIYV 129
DB 68 SVKAMEHASKKLNCEQLQEVPEDPWGRDEANKIAENNDLLWMDYHOKLVQALLTMDTYL 127
QY 130 AQSFIKRIAKRGKLVYDYSARHHLKAVQNA-KDEAKTAKAEEFNKAQTVFEDLNQ 188
DB 128 GQFPDIKSRIAKRGKLVYDYSARHHLKAVQNA-KDEAKTAKAEEFNKAQTVFEDLNQ 187
QY 189 ELLEELPILYNSRIGCVYTFQNIENLSDVFRMSKLNHLYEVMSKLEKQHSNKVYV 248
DB 188 DLQELPILYNSRIGCVYTFQNIENLSDVFRMSKLNHLYEVMSKLEKQHSNKVYV 247
QY 249 KGLSSSRSLVISPPVTRATVSSPLTSPTSLKSESESVSATDLAPDAAGEDN 308
DB 248 K-----AQPRTKSLFSLRRKNSDNAPAKGNKSP-----278
QY 309 SEIKELLEEEIEKEGSEASSEEDDPLPACNGPAQAQPSPTTTERAKSQEVLPSSTPS 368
DB 279 -----SPDGSFAATPEIRVNHE-----PE 298
QY 369 PGGALSPGQPSSTATEVVLRTTASGESEQPKKASIQRTSAPSRPPPPRATASPPS 428
DB 299 PAGGATPGA-----TLPKSPQAEASEVAGGTQPAAGAQEPGETAASEAA 344
QY 429 SGNIPS-----SPTASGGSGPTSPRASLTG-----TASPRTSLEVSPN 467
DB 345 SSSLPAVVVETPPATVNGTVEGSG--AGRLDLPFGFMFKVQAQHDYATATDTDELQKAG 402
QY 468 PEPPEKPVRTPEAKE-----NENIHQNP-E-ELCTS--PTLMTSOV 505
DB 403 DVVLVFPQNPEDQEGWLMGVKESDWNQHKLEKRCGVFPENFTERV 450

```

Db 326 TOPAAGAQPGTSGASEAASSSLPAVVVTFPATVNGTVEGG-----SGAGRLDLP 377

RESULT 5

US-08-435-454-4  
; Sequence 4, Application US/08435454  
; Patent No. 5605830  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; APPLICANT: Sakamuro, Daitoku  
; TITLE OF INVENTION: Murine and Human C-Myc Interacting  
; TITLE OF INVENTION: Protein and Uses Therefor  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,454  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: WST60USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-435-454-4

Query Match 27.7%; Score 792; DB 1; Length 451;  
Best Local Similarity 44.0%; Pred. No. 1.5e-45;  
Matches 184; Conservative 53; Mismatches 119; Indels 62; Gaps 9;

Qy	10	AGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASFYQQAEGHKLKDLKNFLS	69
Db	8	AGKIASNVQKLTFAQEKVLQKLGKADETKDEQFQCVQNFNQLTEGTRLQKDLRTYLA	67
Qy	70	AVKMHSSKRVSETLQEIYSSEWDGHEELKAIWNNDLLWEDYEKLADQAVRTMELY	129
Db	68	SVKAMHEASKKLNCEIQVEYEPDWPGRDEANKIAENNDLLMDYHQKLVQDALLTMDTYL	127
Qy	130	AQFSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAEFEFNKAQTVFEDLNQ	188
Db	128	GQPDIKSRIAKGRKLVYDSARHYESIQAKKDEAKIAKAEELIKAKVFEENV	187
Qy	189	ELLEELPILYNSRIGCVYTFIQNISLRDVFYREMSKLNHNLVYEMSKLEKQHSNKVYV	248
Db	188	DLOEELPSLWNSRVGFYVNTFQSIAGLEENFHKEMSKLNQNLNDVLVGLGKQHSNTFTV	247
Qy	249	KGLSSSSRRSLVSPVVRTATVSPVTSPTSLSKSESSESVATEDLAPDAAQGEDN	308
Db	248	K-----AQRKSKLSRLRRKKNSDNAPAKGNKSPS-----	279
Qy	309	SEIKELLEEEIEKGESEASSSE---EDDPLPACNG-PAQAQPSPTTTERAKSQEVLPS	364
Db	280	-----PPDGSPPATPEIRVNHPEPAGGATPGATLPKSPSPQAEASE--VAGG	325
Qy	365	TTSPG---GALSPSQSSSSATEVLRTRTAS-----EGSEQPKKASIORISAPP	413

RESULT 6

US-08-919-145-6  
; Sequence 6, Application US/08919145  
; Patent No. 5958753  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; TITLE OF INVENTION: Bau, A Binl Interacting Protein, and  
; TITLE OF INVENTION: Uses Therefor  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/919,145  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/025,482  
; FILING DATE: 29-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kodroff, Cathy A.  
; REGISTRATION NUMBER: 33,980  
; REFERENCE/DOCKET NUMBER: WST73AUSA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-919-145-6

Query Match 27.7%; Score 792; DB 2; Length 451;  
Best Local Similarity 44.0%; Pred. No. 1.5e-45;  
Matches 184; Conservative 53; Mismatches 119; Indels 62; Gaps 9;

Qy	10	AGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASFYQQAEGHKLKDLKNFLS	69
Db	8	AGKIASNVQKLTFAQEKVLQKLGKADETKDEQFQCVQNFNQLTEGTRLQKDLRTYLA	67
Qy	70	AVKMHSSKRVSETLQEIYSSEWDGHEELKAIWNNDLLWEDYEKLADQAVRTMELY	129
Db	68	SVKAMHEASKKLNCEIQVEYEPDWPGRDEANKIAENNDLLMDYHQKLVQDALLTMDTYL	127
Qy	130	AQFSEIKERIAKGRKLVYDSARHLEAVQNA-KDEAKTAKAEFEFNKAQTVFEDLNQ	188
Db	128	GQPDIKSRIAKGRKLVYDSARHYESIQAKKDEAKIAKAEELIKAKVFEENV	187
Qy	189	ELLEELPILYNSRIGCVYTFIQNISLRDVFYREMSKLNHNLVYEMSKLEKQHSNKVYV	248
Db	188	DLOEELPSLWNSRVGFYVNTFQSIAGLEENFHKEMSKLNQNLNDVLVGLGKQHSNTFTV	247
Qy	249	KGLSSSSRRSLVSPVVRTATVSPVTSPTSLSKSESSESVATEDLAPDAAQGEDN	308
Db	248	K-----AQRKSKLSRLRRKKNSDNAPAKGNKSPS-----	279
Qy	309	SEIKELLEEEIEKGESEASSSE---EDDPLPACNG-PAQAQPSPTTTERAKSQEVLPS	364
Db	280	-----PPDGSPPATPEIRVNHPEPAGGATPGATLPKSPSPQAEASE--VAGG	325

QY 365 TTPSPG-----GALSPGQPSSTATEVWLRTTAS-----EGSEQPKKRASIQRTSAPP 413  
 Db 326 TOPAAGAEFGTSEASASSLPAAVVVETTPATVNGTVEGG-----SGAGRLDUPP 377

RESULT 7  
 US-09-344-889-6  
 ; Sequence 6, Application US/09344889  
 ; Patent No. 6140465  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prendergast, George C.  
 ; TITLE OF INVENTION: Bau, A Binl Interacting Protein, and  
 ; TITLE OF INVENTION: Uses Therefor  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P.O. Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/344,889  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/919,145  
 ; FILING DATE:  
 ; APPLICATION NUMBER: US 60/025,482  
 ; FILING DATE: 29-AUG-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kodroff, Cathy A.  
 ; REGISTRATION NUMBER: 33,980  
 ; REFERENCE/DOCKET NUMBER: WST73AUSA  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-344-889-6

Query Match 27.7%; Score 792; DB 3; Length 451;  
 Best Local Similarity 44.0%; Pred. No. 1.5e-45;  
 Matches 184; Conservative 53; Mismatches 119; Indels 62; Gaps 9;

QY 10 AGLFAKQVKFSRAQEKVLQKLGKAVETKDERFQSSNFYQQAEGHKLYKDLKNFLS 69  
 Db 8 AGKIASNVQKLTQAEKVLQKLGKAVETKDERFQSSNFYQQAEGHKLYKDLKNFLS 67

QY 70 AVKVMHSSKRVSTLQEIYSSEWDGHEELKAIWNNDLLMEDYEELADQAVRTMEIYV 129  
 Db 68 SVKAMHEASKKLNCEQLQEVTEPDWPGRDEANKIAENNDLLMDYHOKLVDQALLTMDTYL 127

QY 130 AQPSEIKERIAKGRKLVDSARHHLAVQNA-KKDEAKTAKAEFEFNKAQVFEDLNQ 188  
 Db 128 GQFPDIKSIARGRKLVDSARHHLAVQNA-KKDEAKTAKAEFEFNKAQVFEDLNQ 187

QY 189 ELLEELPILYNSRICGVYTFQINSLRDVYFVREMSKLNHLYEVMSKLEKQHSNKVFVV 248  
 Db 188 DLQELPILYNSRICGVYTFQINSLRDVYFVREMSKLNHLYEVMSKLEKQHSNKVFVV 247

QY 249 KGLSSSRSLVISPVRTATVSSPLTSPSTSLKSESESVSATDLAPDAQGEDN 308

Db 248 K-----AQPRKSKLFSRLRRKKNSDNAPAKGNKSPS----- 279  
 QY 309 SEIKELLEEEIEKEGSEASSE---EDDPLACNG-PAQAQSPPTERRAKSOEEVLPS 364  
 Db 280 -----PPDGSFAATPEIRVNHPEPEAGGATFGATLPKSPSQPAEASE--VAGG 325

QY 365 TTPSPG-----GALSPGQPSSTATEVWLRTTAS-----EGSEQPKKRASIQRTSAPP 413  
 Db 326 TOPAAGAEFGTSEASASSLPAAVVVETTPATVNGTVEGG-----SGAGRLDUPP 377

RESULT 8  
 PCT-US96-06231A-4  
 ; Sequence 4, Application PC/TUS9606231A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wistar Institute of Anatomy & Biology  
 ; TITLE OF INVENTION: Murine and Human Box-Dependent  
 ; TITLE OF INVENTION: Myc-Interacting Protein (BIN1) and Uses Therefor  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P O Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/06231A  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/435,454  
 ; FILING DATE: 05-MAY-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bak, Mary E.  
 ; REGISTRATION NUMBER: 31,215  
 ; REFERENCE/DOCKET NUMBER: WST60APCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; PCT-US96-06231A-4

Query Match 27.7%; Score 792; DB 5; Length 451;  
 Best Local Similarity 44.0%; Pred. No. 1.5e-45;  
 Matches 184; Conservative 53; Mismatches 119; Indels 62; Gaps 9;

QY 10 AGLFAKQVKFSRAQEKVLQKLGKAVETKDERFQSSNFYQQAEGHKLYKDLKNFLS 69  
 Db 8 AGKIASNVQKLTQAEKVLQKLGKAVETKDERFQSSNFYQQAEGHKLYKDLKNFLS 67

QY 70 AVKVMHSSKRVSTLQEIYSSEWDGHEELKAIWNNDLLMEDYEELADQAVRTMEIYV 129  
 Db 68 SVKAMHEASKKLNCEQLQEVTEPDWPGRDEANKIAENNDLLMDYHOKLVDQALLTMDTYL 127

QY 130 AQPSEIKERIAKGRKLVDSARHHLAVQNA-KKDEAKTAKAEFEFNKAQVFEDLNQ 188  
 Db 128 GQFPDIKSIARGRKLVDSARHHLAVQNA-KKDEAKTAKAEFEFNKAQVFEDLNQ 187

QY 189 ELLEELPILYNSRICGVYTFQINSLRDVYFVREMSKLNHLYEVMSKLEKQHSNKVFVV 248  
 Db 188 DLQELPILYNSRICGVYTFQINSLRDVYFVREMSKLNHLYEVMSKLEKQHSNKVFVV 247

```

QY 249 KGLSSSSRRSLVSPVTRATVTSPTSTSLKSESESVSATEDLAPDAAGEDN 308
Db 248 K-----AQRKKSKLFSRLRRKKNSDNAPAKGKSPS----- 279
QY 309 SEIKELLEEEIEKGESEASSE---EDDPLPACNG-PAQAQPSPTTTERAKSQEVLPS 364
Db 280 -----PPDGSPPATPEIRVNHEPEFAGGATGATLPKSPSQPAEASE--VAGG 325
QY 365 TTPSPG---GALSPGQSSSATEVLRTRTAS-----EGSEQPKKASTQRTSAPP 413
Db 326 TQPAAGAQBPGETSASEAASSLPAVVVETFPATVNGTVEGG-----SGAGRLDLP 377

RESULT 9
US-08-630-915A-24
; Sequence 24, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: McCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-915A-24

Query Match 23.68; Score 675; DB 4; Length 404;
Best Local Similarity 34.14; Pred. No. 9.9e-38;
Matches 169; Conservative 62; Mismatches 133; Indels 132; Gaps 12;

QY 42 RFEQASNFYQQQAEGHKLKDLNFKLSAVKVMHSESKVSETLQEIYSSSEWDGHEELKA 101
Db 8 RFEQCVQNFNKLTEGTRLQDLRTLYLSVKAHEASKKLNELQSVYEPDPGRDEANK 67

QY 102 IVWNDDLLWEDYEKLADQAVRTWETVVAQFSIKERIAKRGKLVYDSARHLEAVQN 161
Db 68 IAENDDLLWMDYHOKLVQDQALLTMDTVLGQFPDIKSIKRGKLVYDSARHLYESLQT 127

QY 162 A-KXDEAKTAKAEFEFNKAQTVFEDLNQELLELPILYNSRIGCYVTIFQNSLNRDVFY 220

```

```

Db 128 AKKQDEAKIAKAEELIKAKQVFEEEMVDLQEBLPSLMNSRVGYVNTFQSIAGLEENFH 187
QY 221 REMSKLNENLYEVMSKLEKSHSNKVFVVKGLSSSSRRSLVSPVTRATVTSPTSTSP 280
Db 188 KEMSKLNQNTDLVGLKQHGNSITVK----- 216
QY 281 STLKSESESVSATEDLAPDAAGEDNSRIKELLEEBEIEKGESEASSEEDDPLPACN 340
Db 217 -----AQSNDNAPAK-----GNKSPS----- 232
QY 341 GPAQAQPSPTTTERAKSQEVLPSSTTPSPGALSPSQSPSSSATEVVLRTTASEGSEOP 400
Db 233 -PPDGSPPATPEIRVNHE-----PEPAGGATPGA-----TLPKSPSQP 269
QY 401 KKRASIORTSAPPRPPPPPRATASPRSSGNIPS-----SPTASGGSGSPTSPRAS 450
Db 270 AEASEVAGGTQPAAGAQBPGETAASEAASSLPAVVVETFPATVNGTVEGGSG--AGRLD 327
QY 451 LGTG-----TASPRISLEVPSPPEPKPVETPEAKE-----NENIHQNPE 492
Db 328 LPPGFMFVKVQAQHDYATDTDELQKAGDVVLVIPPONPEQDEGMLMGVKSQDWNQHK 387
QY 493 -ELCTS--PTLMTISQV 505
Db 388 LEKCRGVFPENFTERV 403

RESULT 10
US-09-976-594-615
; Sequence 615, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 615
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 2789525CD1
US-09-976-594-615

Query Match 8.5%; Score 243; DB 4; Length 904;
Best Local Similarity 23.4%; Pred. No. 3.4e-08;
Matches 142; Conservative 90; Mismatches 224; Indels 152; Gaps 28;

QY 6 AGGAAGLFAKQVQKFPRAQKVLQKGLKAVETKDERPEQASNFYQQQAEGHKLKDLK 65
Db 3 AGFFRGTSAEQ-DNRFNSNKKQKLLKQL-KFAECLEKKVDMSKVN----- 44
QY 66 NELSVMKVMHSESKVSETLQEIYSSSEWDGHEELKAIWNNDLLWEDYEKLADQAVRTM 125
Db 45 --LEVIKFW--ITKVVITIL-----GFEDDVI-----EFNQLVKNPDS--KMM 85
QY 126 ETVVAQFSIKERIAKRGKLVYDSARHLEAVQNA-----KKDEAKTAKAEFEFNKAQ 181
Db 86 QINLTGFLGNKAREFMGELWPLLSAQENIAGTSPAPLELKKEEIKQRTQIEQ--KLAM 144
QY 182 VFEDLNQELLELPILYNSRIGCYVTIFQNSLNRDVFYREMSKLNENLYEVMSKLEKQ 241
Db 145 KKQDEKDK-----KRDKEKESR-----EKRRSR 170
QY 242 SNKVFVVKGLSSSSRRSLVSPVTRATVTSPTSTSLKSESESVSATEDLAPD 301

```

Db 171 SPR--RRKSRSPRR--SSPVRRERKSHSRSPRH-----RTKSRSP--AP- 214  
 QY 302 AAGGDNSEIKELLBEEIEKEGS--EASSSE-----DPLPACNGPAQAQSPPTT 353  
 Db 215 ----EKKEKTPELPSPVVKPSPVOEATSTSDILKVPKPFIPF--PKPSPKNSKK 267  
 QY 354 AKSQEVLPSSTTPPGGALSPGSPSSSATEVWLRTTASGSEQ---PKKRASIORTS 410  
 Db 268 EKEKTRPRSRSRKSRRTSRSPSHTRPRRRHRSRSRSPRRRSPRRRSPRRRT 327  
 QY 411 APPSRPPPRATASPRP-----SSGNIPSSPTASGGGSPSPRASLTGTAS--PRISLE 463  
 Db 328 PPRMPPPPRHRRSRSPVRRRRSRASLSGSSSSSRSRSPPKPKPTSSPPKTRR 387  
 QY 464 VSNPPPEPK-----PVRTPAEKENENHNQPEP-----LCTSPMTSOVASEPG 510  
 Db 388 LSPSASPPRRHRPPSPATPPKTR--HSPTPQSNTRKSRVSVPGRTSKVTXKG 444  
 QY 511 EAK-----KME-----DKDKNLISADSGDQQLQVSMVPPNNLTAPEQEE 555  
 Db 445 TEKRESPPAPKPKRVKVELSEBEDXGKMAAADSVQORRQYR----RQOQSSSDSGSS 500  
 QY 556 VTSSENPO 563  
 Db 501 SSEDERP 508

RESULT 11  
 US-09-645-456A-13  
 ; Sequence 13, Application US/09645456A  
 ; Patent No. 6562580  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FU, C. Alan  
 ; TITLE OF INVENTION: NOVEL GERMALIN CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND  
 ; FILE REFERENCE: A-68344/RMS/DHR  
 ; CURRENT APPLICATION NUMBER: US/09/645,456A  
 ; CURRENT FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US/09/425,324  
 ; PRIOR FILING DATE: 1999-10-21  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 13  
 ; LENGTH: 1324  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic  
 US-09-645-456A-13

Query Match 7.3%; Score 207.5; DB 4; Length 1324;  
 Best Local Similarity 18.9%; Pred. No. 1.4e-05;  
 Matches 113; Conservative 114; Mismatches 257; Indels 115; Gaps 21;  
 QY 15 KQVQKFSRAQEKVLQKLGKAVETKDERFEQASNFYQQQAEHGK-----LYK 62  
 Db 296 RQVRIQLKDHIDRTKKRGEKDETE---YEYSGSEEEENDSGEPSSILNLPGETLRR 352  
 QY 63 DLKNFLSAVKVNHSSKRVSETLQEIYSEWDGHEELKAIWNNDLLWEDYEE---KLAD 119  
 Db 353 DFLRLQANKERSEALR-----QOLEQQORENEEHKQLLAERQKRIEEOKEORRLEE 407  
 QY 120 QAVRTMEIYVAQFSIKERIAKRGKLVYDSARHLEAVQNAKKADEAKTAAEEFNKA 179  
 Db 408 QORREKELKQOERQRRHVEEQMREERRAEHQEYKQLEQOERQAELOQLQKOE 467  
 QY 180 QTVFDLNGELLELP-----LYNSRIGCVITIFQNISLNRDVFYREMSKLN----- 227  
 Db 468 RDYLSLQHORQOERQVPEKFLYHYKEG--MSPSEKPAWAKEV--EERSRLNRQSSPAMP 523  
 QY 228 HNLVEVMSKLEKQHNKVFVKGLSSSSRRSLVSPVRTATVSPSTSPSTSLSKS 287  
 Db 408 QORREKELKQOERQRRHVEEQMREERRAEHQEYKQLEQOERQAELOQLQKOE 467  
 QY 180 QTVFDLNGELLELP-----LYNSRIGCVITIFQNISLNRDVFYREMSKLN----- 227  
 Db 468 RDYLSLQHORQOERQVPEKFLYHYKEG--MSPSEKPAWAKEV--EERSRLNRQSSPAMP 523  
 QY 228 HNLVEVMSKLEKQHNKVFVKGLSSSSRRSLVSPVRTATVSPSTSPSTSLSKS 287

Db 524 HKVANRISDPNLPSPRSFSISGVQ-----PARTPPMLRPV-DQIPLHVAVKS 571  
 QY 288 ESESVSATEDIAPDAQAQEDNSNSEIKELLEEEIEKEGSEASSSEEDDPLFA-----CN 340  
 Db 572 QCPALTASQSVHEQPTKGLSGFQALNVTSHRVEMPRQNSDFTSENPLTRIEKDRSS 631  
 QY 341 GPAQAQ--PSPPTTERAKSQEVLPSSTTPPGGALSP--SQPSSSATEVWLRTTASEG 396  
 Db 632 WLQOEDIPPKVPQRTTTSIPALARKNSPGNSALGPRGSPQIPRASNPDLRRT----- 685  
 QY 397 SQOPKKRASIORTSAPSRPPPPPRATASPRSSGNIPSS--PTASGGGSPSPRASLTGT 455  
 Db 686 --EPILSPLORTSSGSS-----SSSTFSSQSPGSGSQSGS----- 725  
 QY 456 ASPTSLVSP---NPEPPEKPVRTPEAKENENHNQPEEL-----CTS 497  
 Db 726 -SERTVRANSKSGSPVLPHEPAKVPEESRDTTPSRPADLTALAKELRELRIETNR 784  
 QY 498 PTLMTSOVASPEGEAKMEDKDKNLISADSGEQD--QLQVSMVPPNNLTAPEQEE 555  
 Db 785 PMKVTDYSSSSSESESESEEDGE-----SETHDGTVAVSDIPRLIPTGAPGSE 837

RESULT 12  
 US-09-425-324A-13  
 ; Sequence 13, Application US/09425324A  
 ; Patent No. 6562591  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FU, C. Alan  
 ; TITLE OF INVENTION: NOVEL GERMALIN CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND  
 ; FILE REFERENCE: A-68344/RMS/DHR  
 ; CURRENT APPLICATION NUMBER: US/09/425,324A  
 ; CURRENT FILING DATE: 1999-10-21  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 13  
 ; LENGTH: 1324  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic  
 US-09-425-324A-13

Query Match 7.3%; Score 207.5; DB 4; Length 1324;  
 Best Local Similarity 18.9%; Pred. No. 1.4e-05;  
 Matches 113; Conservative 114; Mismatches 257; Indels 115; Gaps 21;  
 QY 15 KQVQKFSRAQEKVLQKLGKAVETKDERFEQASNFYQQQAEHGK-----LYK 62  
 Db 296 RQVRIQLKDHIDRTKKRGEKDETE---YEYSGSEEEENDSGEPSSILNLPGETLRR 352  
 QY 63 DLKNFLSAVKVNHSSKRVSETLQEIYSEWDGHEELKAIWNNDLLWEDYEE---KLAD 119  
 Db 353 DFLRLQANKERSEALR-----QOLEQQORENEEHKQLLAERQKRIEEOKEORRLEE 407  
 QY 120 QAVRTMEIYVAQFSIKERIAKRGKLVYDSARHLEAVQNAKKADEAKTAAEEFNKA 179  
 Db 408 QORREKELKQOERQRRHVEEQMREERRAEHQEYKQLEQOERQAELOQLQKOE 467  
 QY 180 QTVFDLNGELLELP-----LYNSRIGCVITIFQNISLNRDVFYREMSKLN----- 227  
 Db 468 RDYLSLQHORQOERQVPEKFLYHYKEG--MSPSEKPAWAKEV--EERSRLNRQSSPAMP 523  
 QY 228 HNLVEVMSKLEKQHNKVFVKGLSSSSRRSLVSPVRTATVSPSTSPSTSLSKS 287  
 Db 524 HKVANRISDPNLPSPRSFSISGVQ-----PARTPPMLRPV-DQIPLHVAVKS 571  
 QY 288 ESESVSATEDIAPDAQAQEDNSNSEIKELLEEEIEKEGSEASSSEEDDPLFA-----CN 340  
 Db 572 QCPALTASQSVHEQPTKGLSGFQALNVTSHRVEMPRQNSDFTSENPLTRIEKDRSS 631  
 QY 341 GPAQAQ--PSPPTTERAKSQEVLPSSTTPPGGALSP--SQPSSSATEVWLRTTASEG 396

Db 632 WLRQEDIPPKVQRTTISIPALARKNSPGNSALGRIGSQPIRASNPDLRT----- 685  
 QY 397 SEQPKKASRTQTSAPPSPRRPPRATASPRSSGNIPSS-PTASGGGSPTSRASLTGT 455  
 Db 686 --EPILSPLORTSSGSS-----SSSTPSSQPSQSGSQSGS----- 725  
 QY 456 ASPRTSLVSP---NPEPPKPVRTPEAKENIHQNPPEL-----CTS 497  
 Db 726 -SERTRVANSKSEGSVLPHEPAKVKEPSRDITRSPADLTALAKELRELRIETNR 784  
 QY 498 PTLMTSQVASEPGEAKMEDKEKONKLISADSEGD-QLOVSMVPENNLTAPPEOE 555  
 Db 785 PMKVTYDSSSESESESEEDGE-----SETHDGTVAVSDIPRLIPTGAPGSNEQ 837

## RESULT 13

US-09-645-791-13  
 ; Sequence 13, Application US/09645791  
 ; Patent No. 6569658  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Luo, Ying  
 ; APPLICANT: Fu, Alan C  
 ; APPLICANT: Shen, May  
 ; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS; COMPOSITIONS AN  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; FILE REFERENCE: A-68344-1/RMS/DHR  
 ; CURRENT APPLICATION NUMBER: US/09/645,791  
 ; CURRENT FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 09/425,324  
 ; PRIOR FILING DATE: 1999-10-21  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 13  
 ; LENGTH: 1324  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic  
 US-09-645-791-13

Query Match 7.3%; Score 207.5; DB 4; Length 1324;  
 Best Local Similarity 18.9%; Pred. No. 1.4e-05;  
 Matches 113; Conservative 114; Mismatches 257; Indels 115; Gaps 21;  
 QY 15 KQVQKFSRAQEKVLQKLGKAVETKDERPEQASNFYQQAGHK-----LYK 62  
 Db 296 RQVRIQLKDHIDRTKKRGKDETE---YEYSGSEEEENDSGSPSSILNLPGESTLRR 352  
 QY 63 DLKNFLSAVKVHSESSKRVSETLQEIYSSWDGHEELKAIWVNDLLWEDYEE---KLAD 119  
 Db 353 DFLRLQLANKERSEALR-----QQLQQQRENEEHKQQLAERQKRIEEOKEQRRLEE 407  
 QY 120 QAVTMEIYVAQSEIKERIAKRGKLVYDSARHLEAVQNAKDEAKTAKAEFEFKA 179  
 Db 408 QQRREKLKQERQERQHYEEQMRREERRAEHQYKRLQEBQAEQLRQLKQE 467  
 QY 180 QTVFEDLNQELLEPLI---LYNSRIGCVVTFQNISLNRDVFYREMSKLN----- 227  
 Db 468 RYLVLSLQHQEQRPVEKPLHYKEG--MSPSEKPAWAKEV--EERSRLNRQSSPAMP 523  
 QY 228 HNLVEVMSKLEKQHSNKFVVKVGLSSSSRSRLSVISPPVTRATVSSPLTSPSTLSKS 287  
 Db 524 HKVANRISDPNLPSPRSFSGVQ-----PARTPMLRPV-DPQIPHILVAVKS 571  
 QY 288 ESESVSATDLAPDAAGEDNSEIKELLEBEETEGSEASSEEDDPLPA-----CN 340  
 Db 572 QGPALTASQVHEOPTKGLSGFQALNVTSHRVEMPRQNSDPTSENPPFTTIEKFDRSS 631  
 QY 341 GPAQAQ--PSPTTBRASQSEVLPSSTTPSPGGLSP--SGOPSSSATEVLRTRTASEG 396  
 Db 632 WLRQEDIPPKVQRTTISIPALARKNSPGNSALGRIGSQPIRASNPDLRT----- 685

QY 397 SEQPKKASRTQTSAPPSPRRPPRATASPRSSGNIPSS-PTASGGGSPTSRASLTGT 455  
 Db 686 --EPILSPLORTSSGSS-----SSSTPSSQPSQSGSQSGS----- 725  
 QY 456 ASPRTSLVSP---NPEPPKPVRTPEAKENIHQNPPEL-----CTS 497  
 Db 726 -SERTRVANSKSEGSVLPHEPAKVKEPSRDITRSPADLTALAKELRELRIETNR 784  
 QY 498 PTLMTSQVASEPGEAKMEDKEKONKLISADSEGD-QLOVSMVPENNLTAPPEOE 555  
 Db 785 PMKVTYDSSSESESESEEDGE-----SETHDGTVAVSDIPRLIPTGAPGSNEQ 837

## RESULT 14

US-09-645-456A-11  
 ; Sequence 11, Application US/09645456A  
 ; Patent No. 6562580  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fu, C. Alan  
 ; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS; COMPOSITIONS AN  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; FILE REFERENCE: A-68344/RMS/DHR  
 ; CURRENT APPLICATION NUMBER: US/09/645,456A  
 ; CURRENT FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US/09/425,324  
 ; PRIOR FILING DATE: 1999-10-21  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 11  
 ; LENGTH: 1353  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic  
 US-09-645-456A-11

Query Match 7.2%; Score 205; DB 4; Length 1353;  
 Best Local Similarity 18.7%; Pred. No. 2.1e-05;  
 Matches 119; Conservative 111; Mismatches 247; Indels 158; Gaps 23;  
 QY 15 KQVQKFSRAQEKVLQKLGKAVETKDERPEQASNFYQQAGHK-----LYK 62  
 Db 296 RQVRIQLKDHIDRTKKRGKDETE---YEYSGSEEEENDSGSPSSILNLPGESTLRR 352  
 QY 63 DLKNFLSAVKVHSESSKRVSETLQEIYSSWDGHEELKAIWVNDLLWEDYEEKLADQAV 122  
 Db 353 DFLRLQLANKERSEALR-----QQLQQQRENEEHKQQLI-----AERQKRIEEOKE 400  
 QY 123 RTMEIYVAQSEIKERIAKRGKLVYDSARHLEAVQNAKDEAKTAKAEFEF-----N 177  
 Db 401 QRRRL-----EQQRREKLKQERQERQHYEE--QMRREERRAEHQEYIRQLE 452  
 QY 178 KAQTVFEDLNQELLEPLIYNSRIGCVVTFQNISLNRDV--FYREMSKLNHNLVWMS 235  
 Db 453 EBOQLEILQQLLHEQALLLEYK-----RKLEEQRAERLQRLKQERDYLVSLOH 505  
 QY 236 KLEQKH-----SNKFVVKVGLSSSSRSRLSVISP----- 263  
 Db 506 QRCEQRPVEKPLHYKEGMSSEKPAWAKEVEERSRLNRQSSPAMPKRVANRISDPNLP 565  
 QY 264 -----PVRTATVSSPLTSPSTLSKSESVSATDLAPDAAGEDNSEI 311  
 Db 566 PRSEFSISGVQPAKTPPMLREV-DPQIPHILVAVKSQFALTASQVHEOPTKGLSGFQE 624  
 QY 312 KELLEBEETEGSEASSEEDDPLPA-----CNGPAQAQ--PSPTTBRASQSEVLP 362  
 Db 625 ALNVTSHRVEMPRQNSDPTSENPPFTTIEKFDRSSMLRQEDIPPVKVQRTTISIPALA 684  
 QY 363 SSTTPSPGGLSP--SGOPSSSATEVLRTRTASEGSEQPKKRASIQRTSAPPSPRRPPR 420  
 Db 685 RKNSPGNSALGRIGSQPIRASNPDLRT-----EPILSPLORTSSGSS----- 730  
 QY 421 ATASPRSSGNIPSS-PTASGGGSPTSRASLTGTASPRTSLEVSP-----NPEPPEKPV 475





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:31:50 ; Search time 42.3548 Seconds  
(without alignments)  
86.723 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_23\_35

Perfect score: 61  
Sequence: 1 RAQEKVQLKGLRA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database :

A\_Geneseq\_29Jan04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	108	5	AB06161 Human NS
2	61	100.0	137	4	AB021410 Novel hum
3	61	100.0	252	4	AB062101 Bar domai
4	61	100.0	414	4	AB013005 Novel hum
5	61	100.0	434	2	AA005391 Mouse SH3
6	61	100.0	451	2	AA006602 Human box
7	61	100.0	451	2	AA047295 Human Bin
8	61	100.0	451	2	AA036881 Human Bin
9	61	100.0	451	2	AA04504 Human Bin
10	61	100.0	453	6	AB069631 Human CGD
11	61	100.0	482	6	AB089771 Protein d
12	61	100.0	564	4	AB062100 Human bri
13	61	100.0	572	4	AB021411 Novel hum
14	61	100.0	588	7	AD045951 Rat Prote
15	61	100.0	588	7	AD060848 Rat Prote
16	61	100.0	588	7	AD060840 Rat Prote
17	61	100.0	588	7	AD060844 Rat Prote
18	61	100.0	588	7	AD060852 Rat Prote
19	61	100.0	593	7	AD060842 Human Pro
20	61	100.0	593	7	AD045953 Human Pro
21	61	100.0	593	7	AD060850 Human Pro
22	61	100.0	593	7	AD060846 Human Pro
23	61	100.0	593	7	AD060854 Human Pro
24	61	100.0	594	4	AB020887 Novel hum
25	61	100.0	594	4	AB020324 Novel hum

26	61	100.0	683	7	ADD44887	Add44887 Rat Prote
27	61	100.0	695	5	AAE22091	Aae22091 Human amp
28	61	100.0	695	5	AAU75110	Aau75110 Synaptic
29	61	100.0	695	7	ADD44889	Add44889 Human Pro
30	51	83.6	594	4	ABG22457	Abg22457 Novel hum
31	47	77.0	602	4	ABB63948	Abb63948 Drosophil
32	41	67.2	373	4	AAU03699	Aau03699 Group B S
33	41	67.2	651	5	ABP30317	Abp30317 Streptoco
34	41	67.2	654	5	ABP29732	Abp29732 Streptoco
35	41	67.2	654	5	ABP26469	Abp26469 Streptoco
36	41	67.2	1381	4	ABB57920	Abb57920 Drosophil
37	40	65.6	264	6	ADB07514	Adb07514 Alloiooco
38	40	65.6	297	6	ADB07516	Adb07516 Alloiooco
39	39	63.9	154	5	AAE15583	Aae15583 Mutated G
40	39	63.9	179	5	AAE15580	Aae15580 Mutated G
41	39	63.9	206	5	AAE15579	Aae15579 Mutated G
42	39	63.9	268	5	AAE15309	Aae15309 Glycine m
43	39	63.9	293	5	AAE15308	Aae15308 Glycine m
44	39	63.9	295	5	AAE15582	Aae15582 Alternati
45	39	63.9	506	3	AAE15620	Aae15620 Streptoco
46	39	63.9	510	4	AAU38014	Aau38014 Streptoco
47	39	63.9	510	4	AAU37881	Aau37881 Streptoco
48	39	63.9	510	6	ABU02157	Abu02157 S. pneumo
49	39	63.9	510	6	ABU46197	Abu46197 Protein e
50	38	62.3	80	4	AAU58951	Aau58951 Propionib
51	38	62.3	80	6	ABM55470	Abm55470 Propionib
52	38	62.3	101	4	AA094218	Aa094218 Human pro
53	38	62.3	202	3	ABP75719	Abp75719 Neisseria
54	38	62.3	220	6	ABP78753	Abp78753 N. gonorr
55	38	62.3	243	6	ABP78416	Abp78416 N. gonorr
56	38	62.3	244	4	ABG16068	Abg16068 Novel hum
57	38	62.3	268	3	AAU44888	Aau44888 Human laf
58	38	62.3	331	3	AAU44887	Aau44887 Human laf
59	38	62.3	357	5	AAU47581	Aau47581 Drosophil
60	38	62.3	427	2	AA068607	Aa068607 Dorsalin
61	38	62.3	464	6	ABU43827	Abu43827 Protein e
62	38	62.3	474	3	AA044560	Aa044560 Virulence
63	38	62.3	474	5	ABP54512	Abp54512 Pasteurel
64	38	62.3	481	6	ABM65091	Abm65091 Phototrab
65	38	62.3	490	6	AD032995	Ad032995 Acinetoba
66	38	62.3	509	4	AB071278	Ab071278 Drosophil
67	38	62.3	552	4	AA062176	Aa062176 Human p60
68	38	62.3	646	4	ABG24969	Abg24969 Novel hum
69	38	62.3	944	4	ABG28794	Abg28794 Novel hum
70	38	62.3	1338	4	ABG25447	Abg25447 Novel hum
71	37	60.7	33	4	AB037738	Ab037738 Peptide #
72	37	60.7	209	2	AAU29510	Aau29510 Human lun
73	37	60.7	209	3	AA044458	Aa044458 Human lun
74	37	60.7	209	4	AAE13799	Aae13799 Human lun
75	37	60.7	209	7	ADD66490	Ad066490 Human lun
76	37	60.7	209	7	AD087744	Ad087744 Human lun
77	37	60.7	378	6	ABU45731	Abu45731 Protein e
78	37	60.7	446	6	AD056740	Ad056740 Human sec
79	37	60.7	446	6	AD040587	Ad040587 Human sec
80	37	60.7	446	6	AD091457	Ad091457 Human sec
81	37	60.7	446	7	ADC74008	Adc74008 Human sec
82	37	60.7	447	2	AAW88693	Aaw88693 Secreted
83	37	60.7	447	4	AB050460	Ab050460 Human sec
84	37	60.7	447	6	AB044717	Ab044717 Novel hum
85	37	60.7	447	7	AB026197	Ab026197 Human pro
86	37	60.7	451	2	AAW63687	Aaw63687 Membrane
87	37	60.7	461	3	AAU66724	Aau66724 Membrane
88	37	60.7	461	3	AAU57941	Aau57941 Human tra
89	37	60.7	461	3	AAU33468	Aau33468 Human PRO
90	37	60.7	461	4	AA065573	Aa065573 Human PRO
91	37	60.7	461	4	AAU93271	Aau93271 Human PRO
92	37	60.7	461	4	AAU93593	Aau93593 Human PRO
93	37	60.7	461	4	AAU12398	Aau12398 Human PRO
94	37	60.7	461	4	AAU38931	Aau38931 Human PRO
95	37	60.7	461	4	AA065247	Aa065247 Human PRO
96	37	60.7	461	5	ABB84897	Abb84897 Human PRO
97	37	60.7	461	5	AB095503	Ab095503 Human ang
98	37	60.7	461	5	ABU58062	Abu58062 Human PRO

99 37 60.7 461 6 ABU59140 Novel hum  
 100 37 60.7 461 6 ABU82652 Human sec  
 101 37 60.7 461 6 AB017842 Novel hum  
 102 37 60.7 461 6 AB060571 Human sec  
 103 37 60.7 461 6 ABU13953 Human PRO  
 104 37 60.7 461 6 ABU81096 Human PRO  
 105 37 60.7 461 6 ABU72538 Novel hum  
 106 37 60.7 461 6 ABU66796 Human PRO  
 107 37 60.7 461 6 ABU59877 Novel sec  
 108 37 60.7 461 6 ABU59287 Human sec  
 109 37 60.7 461 6 AB025984 Human PRO  
 110 37 60.7 461 6 AB025067 Human sec  
 111 37 60.7 461 6 ABU58993 Human sec  
 112 37 60.7 461 6 ABU92371 Novel hum  
 113 37 60.7 461 6 ABU59436 Novel hum  
 114 37 60.7 461 6 ABU67072 Human sec  
 115 37 60.7 461 6 ABU92202 Novel hum  
 116 37 60.7 461 6 ABU10908 Human PRO  
 117 37 60.7 461 6 ABU81660 Novel hum  
 118 37 60.7 461 6 ABU88599 Human sec  
 119 37 60.7 461 6 AB034113 Human PRO  
 120 37 60.7 461 6 ADA45973 Novel hum  
 121 37 60.7 461 6 ADA76404 Human PRO  
 122 37 60.7 461 6 ADA19054 Human PRO  
 123 37 60.7 461 6 ADA61677 Homo sapi  
 124 37 60.7 461 6 ADB19462 Novel hum  
 125 37 60.7 461 6 ADB28003 Human PRO  
 126 37 60.7 461 6 ADA86482 Novel hum  
 127 37 60.7 461 6 ADB16046 Human PRO  
 128 37 60.7 461 6 ADA37812 Human PRO  
 129 37 60.7 461 6 ADA47832 Human PRO  
 130 37 60.7 461 6 ADA21498 Human sec  
 131 37 60.7 461 6 ADA10285 Human sec  
 132 37 60.7 461 6 ADA67627 Human PRO  
 133 37 60.7 461 6 ADB30634 Human PRO  
 134 37 60.7 461 6 ADA85930 Novel hum  
 135 37 60.7 461 6 ADA17829 Human PRO  
 136 37 60.7 461 6 ADA97142 Human PRO  
 137 37 60.7 461 6 ADA79446 Human PRO  
 138 37 60.7 461 6 ADA97585 Novel hum  
 139 37 60.7 461 6 ADB16787 Human PRO  
 140 37 60.7 461 6 ADA27937 Human sec  
 141 37 60.7 461 6 ADA91879 Novel hum  
 142 37 60.7 461 6 ADB14942 Human PRO  
 143 37 60.7 461 6 ADB18903 Novel hum  
 144 37 60.7 461 6 ADA94118 Human PRO  
 145 37 60.7 461 6 ADB20014 Human PRO  
 146 37 60.7 461 6 ADB13326 Human PRO  
 147 37 60.7 461 6 ABO43375 Novel hum  
 148 37 60.7 461 6 ADA94517 Human sec  
 149 37 60.7 461 6 ADA74580 Human PRO  
 150 37 60.7 461 6 ADB24813 Human PRO

## ALIGNMENTS

RESULT 1  
 ABB06161  
 ID ABB06161 standard; protein; 108 AA.

XX ABB06161;  
 AC ABB06161;  
 DT 10-MAY-2002 (first entry)

XX Human NS protein sequence SEQ ID NO:253.  
 DE  
 XX

KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;  
 KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;  
 KW vasotropic; antiarteriosclerotic; antinflammatory; dermatological;  
 KW anorectic; muscular; antinfertility; cardiovascular; anticoagulant;  
 KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiac;  
 KW anticonvulsant; antidiabetic; tranquiliser; antidepressant; antiepileptic;

gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;  
 contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;  
 endometriosis; degenerative disease; multiple sclerosis; psoriasis;  
 rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;  
 inflammation; skin disorder; obesity; muscular dystrophy; AIDS;  
 infertility; cardiovascular disease; coagulation disease; hypertension;  
 ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;  
 diabetes; anxiety; depression; schizophrenia; viral disease; stroke;  
 gastric ulcer; Alzheimer's disease.

Homo sapiens.

WO200206315-A2.

24-JAN-2002.

17-JUL-2001; 2001WO-IL000653.

18-JUL-2000; 2000IL-00137345.

15-DEC-2000; 2000IL-00140354.

(COMP-) COMFUGEN LTD.

Mintz L, Freilich S, Bernstein J;

WPI; 2002-155037/20.

N-PSDB; ABL39815.

One hundred and twenty eight novel nucleic acid sequences, useful for treating and diagnosing e.g. cancer, asthma and Alzheimer's.

Claim 6; Page 288; 290pp; English.

ABL39691 to ABL39818 represent novel human nucleic acid sequences  
 encoding the proteins given in ABB06037 to ABB06164. The novel sequences  
 (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,  
 antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,  
 vasotropic, antiarteriosclerotic, antinflammatory, dermatological,  
 anorectic, muscular, anti-HIV, antinfertility, cardiovascular,  
 anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiac,  
 immunomodulator, anticonvulsant, antidiabetic, tranquiliser, antiulcer,  
 antidepressant, gastrointestinal, antiepileptic, cerebroprotective,  
 nootropic and contraceptive activities. The NS can be used in vaccines,  
 gene therapy and antisense therapy. Nucleic acids, expression vectors and  
 antibodies from the present invention can be used for treating and  
 diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative  
 diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,  
 cataracts, restenosis, atherosclerosis, inflammation, skin disorders,  
 glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular  
 disease, coagulation disease, ischaemia, hypertension, asthma, immune  
 disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,  
 depression, schizophrenia, viral disease, gastric ulcers, stroke,  
 Alzheimer's disease and as a contraceptive

Sequence 108 AA;

Query Match 100.0%; Score 61; DB 5; Length 108;

Best Local Similarity 100.0%; Pred. No. 0.006;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13

DB 23 RAQEKVLQKLGKA 35

RESULT 2  
 ABB21410

ID ABB21410 standard; protein; 137 AA.

XX ABB21410;

XX DT 18-FEB-2002 (first entry)  
 CC

DE Novel human diagnostic protein #21401.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 XX 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX N-PSDB; AAS85597.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51769; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 137 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 0.0077;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVQLKLGKA 13  
 DB 52 RAQEKVQLKLGKA 64  
 RESULT 3  
 ID AAB62101  
 XX AAB62101 standard; protein; 252 AA.  
 XX AAB62101;  
 XX 29-MAY-2001 (first entry)  
 XX Bar domain of Bin1 protein.  
 DE  
 XX  
 KW Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
 KW hyperplastic disease; cytostatic; cell growth regulator; Bin1;  
 XX chromosome 4q22.1.  
 OS Homo sapiens.  
 XX WO200116158-A2.  
 XX 08-MAR-2001.  
 XX 30-AUG-2000; 2000WO-US023723.  
 XX 31-AUG-1999; 99US-0151554P.  
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
 XX Prendergast GC, Ge K;  
 XX WPI; 2001-235087/24.  
 XX New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
 PT useful for regulating cell growth, and for diagnosing or treating  
 PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
 PT or hepatocarcinoma.  
 XX  
 PS Example 3; Fig 2; 62pp; English.  
 XX  
 CC The invention provides a human bridging integrator-2 (Bin2) protein. The  
 CC protein can be expressed by standard recombinant methodology. The Bin2  
 CC proteins or peptides are useful in regulating cell growth, cell survival,  
 CC differentiation, endocytosis and actin organization. These peptides or  
 CC proteins are also useful for diagnosing or treating conditions associated  
 CC with inappropriate expression of Bin2, e.g. liver cancer, breast cancer,  
 CC hepatocarcinoma, myeloid and lymphoid leukemias or blood disorders. The  
 CC peptides, nucleic acid sequences or anti-Bin2 antibodies are useful for  
 CC diagnosing inappropriate expression of Bin2. Bin2 is also useful for  
 CC treating disorders associated with excessive Bin1 levels, e.g. liver,  
 CC colorectal, prostate or breast cancers, epithelia cell cancers, melanoma,  
 CC or hyperplastic disease states. The present sequence represents the Bar  
 CC domain of Bin1 used in homology studies with Bin2 protein  
 XX  
 SQ Sequence 252 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 252;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVQLKLGKA 13  
 DB 24 RAQEKVQLKLGKA 36  
 RESULT 4  
 ID AAB62101  
 XX AAB62101 standard; protein; 414 AA.  
 XX AAB62101;  
 XX 13-FEB-2002 (first entry)  
 XX Novel human diagnostic protein #12996.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX

PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 PA (HYSE-) HYSEQ INC.  
 PI Dmanac RT, Liu C, Tang YT;  
 PR N-PSDB; AAS77192.  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS77192.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 20; SEQ ID NO 43364; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 414 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGA 13  
 DB 23 RAQEKVLQKLGA 35  
 RESULT 5  
 AAW05391  
 ID AAW05391 standard; protein; 434 AA.  
 AC AAW05391;  
 XX 18-FEB-1998 (first entry)  
 DT Mouse SH3p9 protein.  
 DE Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
 KW cellular signalling element; cellular structural element; malignancy;  
 KW protein identification; functional domain; protein screening;  
 KW cellular signal transduction process.  
 XX Mus musculus.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 433  
 FT /note= "encoded by CTA"  
 XX WO9631625-A1.  
 XX

PD 10-OCT-1996.  
 XX 04-APR-1996; 96WO-US004454.  
 XX 07-APR-1995; 95US-00417872.  
 PR 03-APR-1996; 96US-00630915.  
 XX (CYTO-) CYTOGEN CORP.  
 PA (UYNC-) UNIV NORTH CAROLINA.  
 XX Sparks AB, Hoffman N, Kay BK, Fowlkes DM, McConnell SJ;  
 PI WPI; 1996-465045/46.  
 DR N-PSDB; AAT39791.  
 XX Identifying polypeptide(s) having specific functional domain (esp. SH3  
 PT domain) - comprises detecting selective binding to recognition unit,  
 PT regardless of sequence homology.  
 XX Claim 54; Fig 35; 174pp; English.  
 CC AAW05386-W05403 represent novel human and mouse Src-homology region 3  
 CC (SH3) domain containing proteins that can be used in the method of the  
 CC invention. SH3 domain containing proteins play a role in signalling and  
 CC structural elements of cells. The method of the invention is for  
 CC identifying polypeptides containing functional domains of interest  
 CC (especially SH3 domains). The method comprises contacting a multivalent  
 CC recognition unit (RU) complex with a number of peptides and identifying  
 CC polypeptides having a selective binding affinity for the RU complex. The  
 CC method is based on functional similarities and does not rely on sequence  
 CC similarities. Prior methods only gave limited success for identifying  
 CC proteins which contain an SH3 domain due to the minimal sequence homology  
 CC among known SH3 proteins. It has been found that small peptide RUS in  
 CC multivalent form have reduced specificity for a given functional domain  
 CC compared to monomer RUS. Multivalent RU complexes are particularly suited  
 CC to screening for polypeptides containing functional domains that are  
 CC similar to, but not identical in sequence to, the original target  
 CC functional domain. The new method enables proteins having a common  
 CC function to be identified. Identification of novel SH3 proteins will be  
 CC useful for a better understanding of cell growth, malignancy, signal  
 CC transduction processes, etc. New candidate drugs can be identified, and  
 CC their specificities (e.g. pharmacological activities) can be assessed  
 CC using the method of the invention  
 XX Sequence 434 AA;  
 Query Match 100.0%; Score 61; DB 2; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 0.025;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGA 13  
 DB 24 RAQEKVLQKLGA 36  
 RESULT 6  
 AAW06602  
 ID AAW06602 standard; protein; 451 AA.  
 AC AAW06602;  
 XX 07-FEB-1997 (first entry)  
 DT Human box-dependent myc-interacting protein (BIN1).  
 DE Box-dependent myc-interacting protein; BIN1; MIP-99; oncogene;  
 KW oncoprotein; breast cancer; liver cancer; apoptosis;  
 KW tumour suppressor protein; hyperplasia; diagnosis; therapy.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Domain 1..222  
 FT



XX WPI; 1998-207038/18.  
 DR N-PSDB; AAC68498.  
 XX  
 PT New isolated BIN1-associated UL-specific protein - which acts as a tumour  
 PT suppressor, used to develop products for treating e.g. cancers, hyper-  
 PT plastic disease states or degenerative diseases.  
 XX  
 PS Disclosure; Fig 3; 22pp; English.  
 XX  
 CC The present invention relates to mammalian BIN1-associated UL-specific  
 CC (Bau) protein. The Bau protein is useful in the detection, diagnosis and  
 CC treatment of cancers or other disorders associated with inappropriate  
 CC BIN1 levels and/or deregulation, deficiency or amplification of the c-Myc  
 CC oncogenes. The proteins are also useful as antigens for the development  
 CC of anti-Bau antisera and antibodies to Bau, or to a desired fragment of  
 CC the Bau protein, as diagnostic reagents, in gene therapy, and in  
 CC screening and developing chemical compounds or proteins which may be used  
 CC for the treatment of cancers characterized by Bau or BIN1, which regulate  
 CC inappropriate MYC levels  
 XX  
 SQ Sequence 451 AA;  
 Query Match 100.0%; Score 61; DB 2; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGKA 13  
 DB 21 RAQEKVLQKLGKA 33  
 RESULT 9  
 AAW94504  
 ID AAW94504 standard; protein; 451 AA.  
 AC AAW94504;  
 DT 22-APR-1999 (first entry)  
 DE Human Bin1 protein.  
 XX  
 KW Bin1; brain-specific; box-dependent myc-interacting protein; cancer;  
 KW diagnosis; hyperplastic disease; tumour suppressor; gene therapy;  
 KW benign prostatic hypertrophy; neurodegeneration.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO9855151-A1.  
 XX  
 PD 10-DEC-1998.  
 XX  
 PF 04-JUN-1998; 98WO-US011647.  
 XX  
 PR 06-JUN-1997; 97US-00870126.  
 XX  
 PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
 XX  
 PI Prendergast GC, Sakamuro D;  
 XX  
 DR WPI; 1999-059881/05.  
 DR N-PSDB; AAX16322.  
 XX  
 PT New splice variant exons of Box-dependent myc-interacting polypeptide -  
 PT associated with loss of tumour suppressor activity, used for diagnosis of  
 PT cancer and in gene therapy.  
 XX  
 PS Example 1; Page 81-82; 133pp; English.  
 XX  
 CC The present invention describes Bin1 (Box-dependent myc-interacting  
 CC protein 1) brain-specific alternative splice variants exon 12A, B, C and  
 CC D. Bin1 specific antibodies (when labelled) are used to detect cancers or  
 CC other hyperplastic conditions (e.g. benign prostatic hypertrophy)

CC associated with a deficit of normal Bin1 and/or aberrant forms of Bin1.  
 CC The same diseases can also be diagnosed at the nucleic acid level using  
 CC fragments of Bin1 nucleotide sequences in standard amplification and/or  
 CC hybridisation assays. Cancers that can be detected are carcinomas and  
 CC epithelial cell tumours, specifically of prostate, liver and colon/  
 CC rectum, also melanoma. Bin1 nucleotide sequences, when included in a  
 CC vector, and Bin1 proteins may also be used to treat these diseases, and  
 CC also degenerative conditions such as neurodegeneration. Bin1 proteins may  
 CC also be used to raise antibodies, and the nucleotide sequences can be  
 CC used to express the corresponding proteins. Also anti-idiotypic  
 CC antibodies can be used therapeutically, and more generally similar  
 CC materials can be used to treat or diagnose any condition involving  
 CC deregulation, defect or amplification of the c-myc oncogene. Bin1  
 CC nucleotide sequence, proteins and antibodies are also useful to screen  
 CC for agents that may be used to treat Bin1-related cancers. The present  
 CC sequence represents human Bin1 protein  
 XX  
 SQ Sequence 451 AA;  
 Query Match 100.0%; Score 61; DB 2; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGKA 13  
 DB 21 RAQEKVLQKLGKA 33  
 RESULT 10  
 ABR69631  
 ID ABR69631 standard; protein; 453 AA.  
 AC ABR69631;  
 DT 12-AUG-2003 (first entry)  
 DE Human CGDD-31 protein.  
 XX  
 KW Human; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
 KW neuroprotective; cerebroprotective; anti-HIV; anti-allergic;  
 KW antiinflammatory; gynaecological; cancer; atherosclerosis; epilepsy;  
 KW Huntington's disease; stroke; AIDS; allergy; placenta; reproductive;  
 KW CGDD; cell growth; cell differentiation; cell death.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2003027263-A2.  
 XX  
 PD 03-APR-2003.  
 XX  
 PF 26-SEP-2002; 2002WO-US031095.  
 XX  
 PR 28-SEP-2001; 2001US-0326399P.  
 PR 05-OCT-2001; 2001US-0327380P.  
 PR 05-OCT-2001; 2001US-0328186P.  
 PR 12-OCT-2001; 2001US-0329690P.  
 PR 26-OCT-2001; 2001US-0345384P.  
 PR 26-OCT-2001; 2001US-0348165P.  
 PR 02-NOV-2001; 2001US-0350213P.  
 PR 09-NOV-2001; 2001US-0344518P.  
 PR 09-NOV-2001; 2001US-0345143P.  
 PR 16-NOV-2001; 2001US-0332375P.  
 PR 03-DEC-2001; 2001US-0336908P.  
 PR 07-DEC-2001; 2001US-0340747P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Chawla NK;  
 PI Elliott VS, Emerling BN, Gandhi AR, Gietzen KJ, Gorvad AE;  
 PI Griffin JA, Hafalia AJA, Ison CH, Kable AE, Kalafus DP;  
 PI Lehr-Mason PM, Lu DAM, Marquis JP, Nguyen DB, Runkumar J;  
 PI Richardson TW, Sapperstein SK, Swarnakar A, Tang YT, Tran UK;  
 PI Warren BA, Xu Y, Yao MG, Yue H, Yue H;

XX WPI; 2003-421159/39.  
 DR N-PSDB; ACC90608.  
 XX  
 PT New human proteins associated with cell growth, differentiation, and  
 PT death (CGDD), useful for diagnosing, treating and preventing diseases or  
 PT conditions associated with the aberrant CGDD expression e.g. cancer,  
 PT AIDS, or epilepsy.  
 XX  
 PS Claim 1; Page 281-282; 350pp; English.  
 XX  
 CC The invention relates to an isolated polypeptide associated with cell  
 CC growth, differentiation and death (CGDD). Also disclosed are the  
 CC polynucleotides encoding the polypeptides. The polypeptides and  
 CC polynucleotides are useful in diagnosing, treating and preventing  
 CC diseases or conditions associated with the decreased expression or over  
 CC expression of CGDD. Such diseases include cell proliferative (e.g.  
 CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's  
 CC disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and  
 CC reproductive disorders, or disorders of the placenta. They are also  
 CC useful in assessing the effects of exogenous compounds on the expression  
 CC of nucleic acid and amino acid sequences of CGDD. The CGDD or its  
 CC fragments are useful in screening compounds for effectiveness as an  
 CC agonist or antagonist of the polypeptides, or in altering the expression  
 CC of the target polynucleotide and compounds that specifically bind to or  
 CC modulate the activity of the polypeptide. Microarrays consisting  
 CC polynucleotides of the invention are useful in monitoring or measuring  
 CC protein-protein interactions, drug-target interactions, and gene  
 CC expression profiles. Sequences given in records AAR69601-AAR69657  
 CC represent CGDD polypeptides of the invention  
 XX  
 SQ Sequence 453 AA;  
 Query Match 100.0%; Score 61; DB 6; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGA 13  
 Db 24 RAQEKVLQKLGA 36  
 RESULT 11  
 ABUS9771  
 ID ABUS9771 standard; protein; 482 AA.  
 XX  
 AC ABUS9771;  
 XX  
 DT 10-JUL-2003 (first entry)  
 XX  
 DE Protein differentially expressed in cardiovascular disease #65.  
 XX  
 KW Cardiovascular disease; arteriosclerosis; ischaemia; angina pectoris;  
 KW myocardial infarction; cardiast; antiarteriosclerotic; antianginal;  
 KW gene therapy; differential gene expression.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003031650-A2.  
 PN  
 XX 17-APR-2003.  
 XX  
 PF 02-OCT-2002; 2002WO-EP011034.  
 XX  
 PR 08-OCT-2001; 2001GB-00024145.  
 XX  
 XX (FARB ) BAYER AG.  
 PA  
 XX Munnes M, Gehrman M, Wick M, Schmitz G;  
 PI  
 XX WPI; 2003-403108/38.  
 DR  
 DR N-PSDB; ACA89944.  
 XX

PT Predicting, diagnosing or prognosing a cardiovascular disease, e.g.  
 PT angina, ischemia, myocardial infarction or arteriosclerosis by detection  
 PT of a polynucleotide in a biological sample comprises detecting a  
 PT hybridization complex.  
 XX  
 PS Claim 3; Page 433-435; 454pp; English.  
 XX  
 CC The invention describes a method of predicting, diagnosing or prognosing  
 CC a cardiovascular disease by detection of a polynucleotide in a biological  
 CC sample comprises hybridising at least one of the polynucleotide to a  
 CC nucleic acid material of a biological sample, thus forming a  
 CC hybridisation complex, and detecting the hybridisation complex. The  
 CC polynucleotides, polypeptides, antisense molecule, antibody and reagent  
 CC are useful for preparing compositions for preventing, predicting or  
 CC diagnosing, or a medicament for treating a cardiovascular disease, e.g.  
 CC arteriosclerosis, ischaemia, angina pectoris, or myocardial infarction.  
 CC This sequence represents a protein identified in the invention a being  
 CC differentially expressed in individuals with cardiovascular disease  
 XX  
 SQ Sequence 482 AA;  
 Query Match 100.0%; Score 61; DB 6; Length 482;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGA 13  
 Db 24 RAQEKVLQKLGA 36  
 RESULT 12  
 AAB62100  
 ID AAB62100 standard; protein; 564 AA.  
 XX  
 AC AAB62100;  
 XX  
 DT 29-MAY-2001 (first entry)  
 XX  
 DE Human bridging integrator-2 (Bin2) protein.  
 XX  
 KW Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
 KW hyperplastic disease; cytostatic; cell growth regulator;  
 KW chromosome 4q22.1.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 23..35  
 FT Peptide /note= "specifically claimed fragment"  
 FT Peptide 138..155  
 FT /note= "specifically claimed fragment"  
 XX  
 PN WO200116158-A2.  
 XX  
 PD 08-MAR-2001.  
 XX  
 PF 30-AUG-2000; 2000WO-US023723.  
 XX  
 PR 31-AUG-1999; 99US-0151554P.  
 XX  
 PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
 XX  
 PI Prendergast GC, Ge K;  
 XX  
 DR WPI; 2001-235087/24.  
 DR N-PSDB; AAF57268.  
 XX  
 XX New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
 PT useful for regulating cell growth, and for diagnosing or treating  
 PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
 PT or hepatocarcinoma.  
 XX  
 PS Claim 1; Fig 1A-C; 62pp; English.



XX This represents a human bridging integrator-2 (Bin2) protein. The Bin2  
 CC protein can be expressed by standard recombinant methodology. The Bin2  
 CC proteins or peptides are useful in regulating cell growth, cell survival,  
 CC differentiation, endocytosis and actin organization. These peptides or  
 CC proteins are also useful for diagnosing or treating conditions associated  
 CC with inappropriate expression of Bin2, e.g. liver cancer, breast cancer,  
 CC hepatocarcinoma, myeloid and lymphoid leukemias or blood disorders. The  
 CC peptides, nucleic acid sequences or anti-Bin2 antibodies are useful for  
 CC diagnosing inappropriate expression of Bin2. Bin2 is also useful for  
 CC treating disorders associated with excessive Bin1 levels, e.g. liver,  
 CC colorectal, prostate or breast cancers, epithelial cell cancers, melanoma,  
 CC or hyperplastic disease states

XX  
 CC Sequence 564 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 |||||  
 Db 23 RAQEKVLQKLGKA 35

RESULT 13  
 ABG21411  
 ID ABG21411 standard; protein; 572 AA.  
 XX  
 AC ABG21411;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #21402.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Dermanac RT, Liu C, Tang YT;  
 XX  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS85598.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51770; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (I) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pot\_sequences

XX  
 CC Sequence 572 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 572;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 |||||  
 Db 30 RAQEKVLQKLGKA 42

RESULT 14  
 ADD45951  
 ID ADD45951 standard; protein; 588 AA.  
 XX  
 AC ADD45951;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein O08839, SEQ ID NO 11623.  
 XX  
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GEHO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 PI WPI; 2003-268312/26.  
 DR GENBANK; O08839.  
 DR  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PT  
 XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 100.0%; Score 61; DB 7; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 0.035;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGA 13  
 DB 24 RAQEKVLQKLGA 36  
 |||||

RESULT 15  
 ADE60848  
 ID ADE60848 standard; protein; 588 AA.

XX AC ADE60848;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein O08839, SEQ ID NO 6761.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX FN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; O08839.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 100.0%; Score 61; DB 7; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 0.035;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGA 13  
 DB 24 RAQEKVLQKLGA 36  
 |||||

RESULT 16

ADE60840

ID ADE60840 standard; protein; 588 AA.

XX AC ADE60840;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein O08839, SEQ ID NO 6753.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX FN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; O08839.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
 SQ Sequence 588 AA;  
 Query Match 100.0%; Score 61; DB 7; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 0.035;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGA 13  
 |||||  
 Db 24 RAQEKVLQKLGA 36

RESULT 17  
 ADE60844  
 ID ADE60844 standard; protein; 588 AA.

XX AC ADE60844;

DT 29-JAN-2004 (first entry)

DE Rat Protein O08839, SEQ ID NO 6757.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; O08839.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 100.0%; Score 61; DB 7; Length 588;

Best Local Similarity 100.0%; Pred. No. 0.035;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGA 13  
 |||||  
 Db 24 RAQEKVLQKLGA 36

RESULT 18

ADE60852

ID ADE60852 standard; protein; 588 AA.

XX AC ADE60852;

DT 29-JAN-2004 (first entry)

DE Rat Protein O08839, SEQ ID NO 6765.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; O08839.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

SQ Query Match 100.0%; Score 61; DB 7; Length 588;

Best Local Similarity 100.0%; Pred. No. 0.035;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGA 13

DB 24 RAQEKVLQKLGA 36

RESULT 19

AD560842

ID ADE60842 standard; protein; 593 AA.

XX AC ADE60842;

XX 29-JAN-2004 (first entry)

DE Human Protein Q99688, SEQ ID NO 6755.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

OS Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

PT

XX

DR

XX

XX

PT

XX

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

WPI; 2003-269312/26.

GENBANK; Q99688.

XX New composition comprising two or more isolated polypeptides, useful for  
 CC preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 593 AA;

Query Match 100.0%; Score 61; DB 7; Length 593;

Best Local Similarity 100.0%; Pred. No. 0.035;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGA 13

DB 24 RAQEKVLQKLGA 36

RESULT 20

ADD45953

ID ADD45953 standard; protein; 593 AA.

XX AC ADD45953;

XX 29-JAN-2004 (first entry)

DE Human Protein Q99688, SEQ ID NO 11625.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

OS Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.



XX WO2003016475-A2.  
 XX 27-FEB-2003.  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX 14-AUG-2001; 2001US-0312147P.  
 XX 01-NOV-2001; 2001US-0346382P.  
 XX 26-NOV-2001; 2001US-0333347P.  
 XX (GHEO ) GEN HOSPITAL CORP.  
 XX (FARB ) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 XX GENBANK; Q99688.  
 XX New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 XX or human polynucleotides or a polynucleotide which represents a fragment,  
 XX derivative or allelic variation of the nucleic acid sequence. Also  
 XX claimed are a vector comprising the novel polynucleotide, a host cell  
 XX comprising the vector, a method for identifying a nucleotide sequence  
 XX which is differentially regulated in an animal subjected to pain and a  
 XX kit to perform the method, an array, a method for identifying an agent  
 XX that increases or decreases the expression of the polynucleotide sequence  
 XX that is differentially regulated in neuronal tissue of a first animal  
 XX subjected to pain, a method for identifying a compound which regulates  
 XX the expression of a polynucleotide sequence which is differentially  
 XX expressed in an animal subjected to pain, a method for identifying a  
 XX compound that regulates the activity of one or more of the  
 XX polynucleotides, a method for producing a pharmaceutical composition, a  
 XX method for identifying a compound or small molecule that regulates the  
 XX activity in an animal of one or more of the polypeptides given in the  
 XX specification, a method for identifying a compound useful in treating  
 XX pain and a pharmaceutical composition comprising the one or more  
 XX polypeptides or their antibodies. The polynucleotide or the compound that  
 XX modulates its activity is useful for preparing a medicament for treating  
 XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 XX therapy). The sequence presented is a human protein (shown in Table 2 of  
 XX the specification) which is differentially expressed during pain. Note:  
 XX The sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic form directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 593 AA;  
 XX Query Match 100.0%; Score 61; DB 7; Length 593;  
 XX Best Local Similarity 100.0%; Pred. No. 0.035;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX QY 1 RAQEKVLQKLGKA 13  
 XX Db 24 RAQEKVLQKLGKA 36  
 XX RESULT 23  
 XX ADE60854  
 XX ID ADE60854 standard; protein; 593 AA.  
 XX XX  
 XX AC ADE60854;  
 XX XX  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human Protein Q99688, SEQ ID NO 6767.  
 XX XX

KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 XX WO2003016475-A2.  
 XX 27-FEB-2003.  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX 14-AUG-2001; 2001US-0312147P.  
 XX 01-NOV-2001; 2001US-0346382P.  
 XX 26-NOV-2001; 2001US-0333347P.  
 XX (GHEO ) GEN HOSPITAL CORP.  
 XX (FARB ) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 XX GENBANK; Q99688.  
 XX New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 XX or human polynucleotides or a polynucleotide which represents a fragment,  
 XX derivative or allelic variation of the nucleic acid sequence. Also  
 XX claimed are a vector comprising the novel polynucleotide, a host cell  
 XX comprising the vector, a method for identifying a nucleotide sequence  
 XX which is differentially regulated in an animal subjected to pain and a  
 XX kit to perform the method, an array, a method for identifying an agent  
 XX that increases or decreases the expression of the polynucleotide sequence  
 XX that is differentially expressed in neuronal tissue of a first animal  
 XX subjected to pain, a method for identifying a compound which regulates  
 XX the expression of a polynucleotide sequence which is differentially  
 XX expressed in an animal subjected to pain, a method for identifying a  
 XX compound that regulates the activity of one or more of the  
 XX polynucleotides, a method for producing a pharmaceutical composition, a  
 XX method for identifying a compound or small molecule that regulates the  
 XX activity in an animal of one or more of the polypeptides given in the  
 XX specification, a method for identifying a compound useful in treating  
 XX pain and a pharmaceutical composition comprising the one or more  
 XX polypeptides or their antibodies. The polynucleotide or the compound that  
 XX modulates its activity is useful for preparing a medicament for treating  
 XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 XX therapy). The sequence presented is a human protein (shown in Table 2 of  
 XX the specification) which is differentially expressed during pain. Note:  
 XX The sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic form directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 593 AA;  
 XX Query Match 100.0%; Score 61; DB 7; Length 593;  
 XX Best Local Similarity 100.0%; Pred. No. 0.035;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX QY 1 RAQEKVLQKLGKA 13  
 XX Db 24 RAQEKVLQKLGKA 36  
 XX RESULT 24  
 XX ABG20887  
 XX ID ABG20887 standard; protein; 594 AA.  
 XX XX  
 XX AC ABG20887;

XX DT 13-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #20878.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX DR WPI; 2001-639362/73.  
 XX DR N-PSDB; AAS85074.  
 XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX PS Claim 20; SEQ ID NO 51246; 103pp; English.  
 XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC patent did not appear in the printed specification. Note: The sequence data for this  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 594 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 594;  
 Best Local Similarity 100.0%; Pred. NO. 0.035;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGKA 13  
 Db 52 RAQEKVLQKLGKA 64  
 RESULT 25  
 ABG20324  
 ID ABG20324 standard; protein; 594 AA.  
 XX AC ABG20324;  
 XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #20315.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX DR WPI; 2001-639362/73.  
 XX DR N-PSDB; AAS84511.  
 XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX PS Claim 20; SEQ ID NO 50683; 103pp; English.  
 XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC patent did not appear in the printed specification. Note: The sequence data for this  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 594 AA;  
 Query Match 100.0%; Score 61; DB 4; Length 594;  
 Best Local Similarity 100.0%; Pred. NO. 0.035;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGKA 13  
 Db 52 RAQEKVLQKLGKA 64  
 RESULT 26  
 ADD44887  
 ID ADD44887 standard; protein; 583 AA.  
 XX AC ADD44887;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Rat Protein O08838, SEQ ID NO 10318.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 OS Rattus norvegicus.  
 XX WO2003016475-A2.  
 XX 27-FEB-2003.  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX 14-AUG-2001; 2001US-03412147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX (GEO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI: 2003-268312/26.  
 DR GENBANK; C08838.  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX Claim 1; Page; 1017pp; English.  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 683 AA;  
 XX Query Match 100.0%; Score 61; DB 7; Length 683;  
 XX Best Local Similarity 100.0%; Pred. No. 0.04;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGA 13  
 Db 19 RAQEKVLQKLGA 31  
 RESULT 27  
 AA22091  
 ID AA22091 standard; protein; 695 AA.  
 XX  
 AC AA22091;  
 XX Query Match 100.0%; Score 61; DB 7; Length 683;  
 XX Best Local Similarity 100.0%; Pred. No. 0.04;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGA 13  
 Db 19 RAQEKVLQKLGA 31  
 RESULT 28  
 AAU75110  
 ID AAU75110 standard; protein; 695 AA.  
 XX  
 AC AAU75110;  
 XX  
 XX 09-APR-2002 (first entry)  
 XX Synaptic vesicle and endocytosis associated protein, amphiphysin.  
 XX MAPKAP-K3; AP-3 delta; APP-695; amyloid A4; Hsp8; heat shock protein 8;  
 XX L130; NY-REN-58; P38 Alpha; P38 Beta; ERK3; KIAA0934; CDK9;  
 XX cell cycle dependent kinase 9; CLK; C-NAP-1; clathrin heavy chain;  
 XX amphiphysin; Png109; KIAA1106; neurodegenerative disorder;  
 XX non-insulin dependent diabetes mellitus; NIDDM; Alzheimer's disease;  
 XX inflammatory disease; rheumatoid arthritis; inflammatory bowel disorder;  
 XX atherosclerosis; cardiac hypertrophy; hypoxic brain injury;  
 XX yeast two-hybrid; signal transduction pathway; human; endocytosis;  
 XX synaptic vesicle; mitogen activated protein kinase.  
 XX Homo sapiens.  
 XX

XX 25-JUL-2002 (first entry)  
 XX Human amphiphysin-1 protein.  
 XX Cellular senescence; amphiphysin-1 protein; caveolin-1 protein; human;  
 KW gene therapy.  
 XX Homo sapiens.  
 XX WO200221140-A1.  
 XX 14-MAR-2002.  
 XX 06-JUL-2001; 2001WO-KR001159.  
 XX 08-SEP-2000; 2000KR-00053341.  
 PR 08-SEP-2000; 2000KR-00053342.  
 XX (META-) METABOLIC ENG LAB CO LTD.  
 PA Park S, Park W, Park J, Cho K, Kim D;  
 XX WPI: 2002-362263/39.  
 DR N-PSDB; RAD35148.  
 XX Modulating cellular senescence in patient involves administering protein  
 PT involved in cellular senescence e.g., amphiphysin protein or caveolin  
 PT protein or polynucleotide encoding the proteins.  
 XX Claim 23; Page 92-95; 103pp; English.  
 XX The invention relates to a method of modulating cellular senescence in a  
 CC patient. The method involves administering protein involved in cellular  
 CC senescence e.g. amphiphysin-1 protein or caveolin-1 protein or  
 CC polynucleotides encoding such proteins. The method is useful for  
 CC detecting and modulating cellular senescence in a mammalian cell. It is  
 CC also used in gene therapy. The present sequence is human amphiphysin-1  
 CC protein  
 XX Sequence 695 AA;  
 XX Query Match 100.0%; Score 61; DB 5; Length 695;  
 XX Best Local Similarity 100.0%; Pred. No. 0.041;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGA 13  
 Db 19 RAQEKVLQKLGA 31  
 RESULT 28  
 AAU75110  
 ID AAU75110 standard; protein; 695 AA.  
 XX  
 AC AAU75110;  
 XX  
 XX 09-APR-2002 (first entry)  
 XX Synaptic vesicle and endocytosis associated protein, amphiphysin.  
 XX MAPKAP-K3; AP-3 delta; APP-695; amyloid A4; Hsp8; heat shock protein 8;  
 XX L130; NY-REN-58; P38 Alpha; P38 Beta; ERK3; KIAA0934; CDK9;  
 XX cell cycle dependent kinase 9; CLK; C-NAP-1; clathrin heavy chain;  
 XX amphiphysin; Png109; KIAA1106; neurodegenerative disorder;  
 XX non-insulin dependent diabetes mellitus; NIDDM; Alzheimer's disease;  
 XX inflammatory disease; rheumatoid arthritis; inflammatory bowel disorder;  
 XX atherosclerosis; cardiac hypertrophy; hypoxic brain injury;  
 XX yeast two-hybrid; signal transduction pathway; human; endocytosis;  
 XX synaptic vesicle; mitogen activated protein kinase.  
 XX Homo sapiens.  
 XX



PH Key Location/Qualifiers  
 FT 93..273  
 FT /note= "This region binds the centrosomal Nek-2  
 FT associated protein 1 (C-NAP1) (see ABK1313), the bait  
 FT protein in a yeast two-hybrid assay, producing a complex  
 FT claimed in claim 1"  
 XX  
 XX W0200198524-A2.  
 XX  
 XX 27-DEC-2001.  
 XX  
 XX 21-JUN-2001; 2001WO-US019762.  
 XX  
 XX 22-JUN-2000; 2000US-0213245P.  
 XX  
 XX (MYRI-) MYRIAD GENETICS INC.  
 XX  
 XX Heichman K, Bartel PL;  
 XX WPI; 2002-122287/16.  
 XX  
 XX New protein complexes comprising protein-protein interactions (e.g.  
 PT MAPKAP-K3/AP-3 delta or C-NAP-1/Claithrin HC), useful for diagnosing  
 PT physiological generative disorders or screening drugs for these diseases.  
 XX  
 XX Example 11; Page: 60pp; English.  
 XX  
 CC The invention describes an isolated protein complex, comprising two  
 CC proteins. The protein complex comprises: protein kinase MAPKAP-K3 and AP-  
 CC 3 adaptor complex delta sub-unit; MAPKAP-K3 and amyloid A4 precursor  
 CC protein, APP-695; MAPKAP-K3 and heat shock protein (Hsp) 8; leucine rich  
 CC Li30 and NY-REN-59; P38 Alpha and P38 Beta; protein kinase ERK3 and  
 CC KIAA0934 (unknown function); ERK3 and cell cycle dependent kinase (CDK) 9  
 CC ; ERK3 and protein kinase CLK; C-NAP-1 and Claithrin heavy chain; C-NAP-1  
 CC and Amphiphysin; C-NAP-1 and novel protein. PMS109 or C-NAP-1 and KIAA1106  
 CC (unknown function) interactions. The protein complexes are useful for  
 CC diagnosing physiological generative disorders, drug screening for agents  
 CC that modulate the interaction of the proteins (thus identify drug  
 CC targets), and identifying additional proteins in the pathway common to  
 CC the proteins. These physiological disorders include non-insulin dependent  
 CC diabetes mellitus (NIDDM), neurodegenerative disorders (e.g. Alzheimer's  
 CC disease), inflammatory diseases (e.g. rheumatoid arthritis and  
 CC inflammatory bowel disorder) and other human disease such as  
 CC atherosclerosis, cardiac hypertrophy and hypoxic brain injury. This  
 CC sequence represents the synaptic vesicle and endocytosis associated  
 CC protein amphiphysin, residues 93-273 of which binds to the bait protein  
 CC centrosomal Nek-2 associated protein 1 (C-NAP1) (see ABK1313) in a yeast  
 CC two-hybrid assay for determining components of signal transduction  
 CC pathways and forms an interaction claimed in claim 1 of the invention.  
 CC Note: This sequence does not appear in the specification but has been  
 CC obtained from a reference given in the invention  
 XX  
 XX Sequence 695 AA;  
 SQ  
 Query Match 100.0%; Score 61; DB 5; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 0.041;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLKGA 13  
 Db 19 RAQEKVLQKLKGA 31  
 RESULT 29  
 ID ADD44889  
 XX ADD44889 standard; protein; 695 AA.  
 XX  
 XX ADD44889;  
 XX  
 XX 29-JAN-2004 (first entry)  
 XX  
 XX Human Protein P49419, SEQ ID NO 10320.  
 DE  
 XX

KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 OS  
 XX W02003016475-A2.  
 PN  
 XX 27-FEB-2003.  
 PD  
 XX  
 XX 14-AUG-2002; 2002WO-US025765.  
 PF  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR  
 XX 01-NOV-2001; 2001US-0346382P.  
 PR  
 XX 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 PA  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 PI WPI; 2003-269312/26.  
 DR GENBANK; P49418.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PT  
 XX Claim 1; Page: 1017pp; English.  
 PS  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 695 AA;  
 SQ  
 Query Match 100.0%; Score 61; DB 7; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 0.041;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLKGA 13  
 Db 19 RAQEKVLQKLKGA 31  
 RESULT 30  
 ID ABG22457  
 XX ABG22457 standard; protein; 594 AA.  
 XX  
 XX AC ABG22457;

XX DT 18-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #22448.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX DR WPI; 2001-639362/73.  
 XX DR N-PSDB; AAS86644.  
 XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX PS Claim 20; SEQ ID NO 52816; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 594 AA;  
 Query Match 83.6%; Score 51; DB 4; Length 594;  
 Best Local Similarity 84.6%; Pred. No. 2;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGKA 13  
 DB 52 RAQEKALXKLGKA 64  
 RESULT 31  
 ABB63948  
 ID ABB63948 standard; protein; 602 AA.  
 AC ABB63948;  
 XX 26-MAR-2002 (first entry)  
 DT

XX DE Drosophila melanogaster polypeptide SEQ ID NO 18636.  
 XX KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical.  
 XX OS Drosophila melanogaster.  
 XX PN WO200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US009231.  
 XX PR 23-MAR-2000; 2000US-0191637P.  
 XX PR 11-JUL-2000; 2000US-00614150.  
 XX PA (PEKE) PE CORP NY.  
 XX PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX DR WPI; 2001-656860/75.  
 XX DR N-PSDB; ABL08051.  
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX PS Disclosure; SEQ ID NO 18636; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 602 AA;  
 Query Match 77.0%; Score 47; DB 4; Length 602;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGK 12  
 DB 19 RAKSKILQNLQK 30  
 RESULT 32  
 AAU03699  
 ID AAU03699 standard; protein; 373 AA.  
 AC AAU03699;  
 XX 12-SEP-2001 (first entry)  
 DT Group B Streptococcus antigenic protein, ID-176.  
 DE Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;  
 KW meningitis; neonate; antigenic; vaccine; infection; genital tract;  
 KW capsid polysaccharide vaccination.  
 XX Streptococcus agalactiae.  
 OS WO200132882-A2.  
 XX 10-MAY-2001.  
 PD 07-SEP-2000; 2000WO-GB003437.  
 PF

XX PR 07-SEP-1999; 99GB-00021125.  
 XX PA (MICR-) MICROBIAL TECHNIQS LTD.  
 XX PI Le Page RWF, Wells JM, Hanniffy SB;  
 XX PS WPI; 2001-316444/33.  
 XX DR N-PSDB; AAS07116.  
 XX PT New polypeptides derived from Streptococcus agalactiae are useful to  
 XX PT provide detection of, and vaccination against, Group B Streptococcus  
 XX PT infections, particularly to prevent infection in neonates.  
 XX PS Claim 1; Fig 1; 178pp; English.  
 XX CC AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus  
 CC agalactiae) amino acid sequences of the invention. S. agalactiae is an  
 CC encapsulated bacterium which is a major pathogen of humans causing sepsis  
 CC and meningitis in neonates as well as adults. The S. agalactiae antigenic  
 CC polypeptides are used to vaccinate against Group B Streptococcus  
 CC infections, particularly to prevent infection in new born children  
 CC arising from the maternal genital tract. An immunogenic composition is  
 CC useful in the preparation of a medicament for the treatment or  
 CC prophylaxis of Group B Streptococcus infection. The invention does not  
 CC have the disadvantages of varied response rate associated with prior art  
 CC capsid polysaccharide vaccination against Group B Streptococcus  
 XX SQ Sequence 373 AA;  
 Query Match 67.2%; Score 41; DB 4; Length 373;  
 Best Local Similarity 63.6%; Pred. No. 69;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 QEKVLQKLGKA 13  
 DB 337 EEKILKKLGKS 347  
 RESULT 33  
 ABP30317  
 ID ABP30317 standard; protein; 651 AA.  
 XX AC ABP30317;  
 XX DT 02-JUL-2002 (first entry)  
 XX DE Streptococcus polypeptide SEQ ID NO 9810.  
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus agalactiae.  
 XX PN WO200234771-A2.  
 XX PD 02-MAY-2002.  
 XX PF 29-OCT-2001; 2001WO-GB004789.  
 XX PR 27-OCT-2000; 2000GB-00026333.  
 XX PR 24-NOV-2000; 2000GB-00028727.  
 XX PR 07-MAR-2001; 2001GB-00005640.  
 XX PA (CHIR-) CHIRON SPA.  
 XX PA (GENO-) INST GENOMIC RES.  
 XX PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;  
 XX PI Tettelin H;  
 XX DR WPI; 2002-352536/38.  
 XX DR N-PSDB; ABN70948.  
 XX PT New Streptococcus protein for the treatment or prevention of infection or  
 DR disease caused by Streptococcus bacteria, such as meningitis, and for

XX PT New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for  
 XX PT detecting a compound that binds to the protein.  
 XX PS Claim 1; Page 4109; 4525pp; English.  
 XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABM66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins  
 XX SQ Sequence 651 AA;  
 Query Match 67.2%; Score 41; DB 5; Length 651;  
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 QEKVLQKLGKA 13  
 DB 334 EEKILKKLGKS 344  
 RESULT 34  
 ABP29732  
 ID ABP29732 standard; protein; 654 AA.  
 XX AC ABP29732;  
 XX DT 02-JUL-2002 (first entry)  
 XX DE Streptococcus polypeptide SEQ ID NO 8640.  
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus agalactiae.  
 XX PN WO200234771-A2.  
 XX PD 02-MAY-2002.  
 XX PF 29-OCT-2001; 2001WO-GB004789.  
 XX PR 27-OCT-2000; 2000GB-00026333.  
 XX PR 24-NOV-2000; 2000GB-00028727.  
 XX PR 07-MAR-2001; 2001GB-00005640.  
 XX PA (CHIR-) CHIRON SPA.  
 XX PA (GENO-) INST GENOMIC RES.  
 XX PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;  
 XX PI Tettelin H;  
 XX DR WPI; 2002-352536/38.  
 XX DR N-PSDB; ABN70363.  
 XX PT New Streptococcus protein for the treatment or prevention of infection or  
 PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX Claim 1; Page 3977; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A Streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins

XX Sequence 654 AA;

Query Match 67.2%; Score 41; DB 5; Length 654;  
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGA 13

DB 337 EEKILKLGKS 347

RESULT 35

ABP26469

ID ABP26469 standard; protein; 654 AA.

XX AC ABP26469;

XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 2114.

XX Streptococcus; GAS; GBS; group B Streptococcus; Streptococcus agalactiae;  
 XX group A Streptococcus; Streptococcus pyogenes; antibacterial;  
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

XX 24-NOV-2000; 2000GB-00028727.

XX 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Maignani V, Margarit Y RosI, Grandi G, Fraser C;

XX Tettelin H;

XX WPI; 2002-352536/38.

XX DR N-PSDB; ABN67100.

XX New Streptococcus protein for the treatment or prevention of infection or

XX disease caused by Streptococcus bacteria, such as meningitis, and for  
 XX detecting a compound that binds to the protein.

XX Claim 1; Page 3364; 4525pp; English.

XX

CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A Streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins

XX Sequence 654 AA;

Query Match 67.2%; Score 41; DB 5; Length 654;  
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGA 13

DB 337 EEKILKLGKS 347

RESULT 36

ABB57920

ID ABB57920 standard; protein; 1381 AA.

XX AC ABB57920;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 552.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-556860/75.

XX N-PSDB; ABL02023.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions.

XX Disclosure; SEQ ID NO 552; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pt\_sequences  
 XX  
 SQ Sequence 1381 AA;  
 Query Match 67.2%; Score 41; DB 4; Length 1381;  
 Best Local Similarity 81.8%; Pred. NO. 2.7e+02;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 QEKVLQKLGA 13  
 | | | | |  
 Db 663 QSKELQKLGA 673  
 RESULT 37  
 ADB07514  
 ID ADB07514 standard; protein; 264 AA.  
 AC ADB07514;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Alloiooccus otitis antigenic protein SEQ ID NO:1454.  
 XX  
 KW Alloiooccus otitis; antigenic protein; immunogenic; immunisation;  
 XX gene therapy; Gram-positive bacterium; infection.  
 XX  
 OS Alloiooccus otitis.  
 XX  
 FN WO2003048304-A2.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PF 25-NOV-2002; 2002WO-US036123.  
 XX  
 PR 29-NOV-2001; 2001US-0333777P.  
 PR 18-NOV-2002; 2002US-0426742P.  
 XX  
 PA (AMHP) WYETH HOLDINGS CORP.  
 XX  
 PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;  
 XX  
 DR WPI; 2003-505284/47.  
 DR N-PSDB; ADB07513.  
 XX  
 PT New Alloiooccus otitis polynucleotides and polypeptides, useful for  
 PT treating and diagnosing diseases, drug screening assays and monitoring of  
 PT effects during drug clinical trials.  
 XX  
 PS Claim 33; SEQ ID NO 1454; 1019pp; English.  
 XX  
 CC The present invention describes an isolated polynucleotide (1) of  
 CC Alloiooccus otitis genomic DNA, which encodes an antigenic protein.  
 CC Alloiooccus otitis is a Gram-positive bacterium. Also described: (1)  
 CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an  
 CC expression vector comprising the novel isolated polynucleotide (1); its  
 CC complement, degenerate variant or fragment; (3) a genetically engineered  
 CC host cell, transfected, transformed or infected with the vector of (2);  
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
 CC composition comprising the polypeptide, its complement, biological  
 CC equivalent or fragment, or the polynucleotide that is comprised in the  
 CC expression vector; (6) a pharmaceutical composition comprising the  
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
 CC of the polypeptides of (1); their biological equivalent or fragment; (8)  
 CC immunising against Alloiooccus otitis by administering to a host the  
 CC immunogenic composition; (9) detecting and/or identifying Alloiooccus  
 CC otitis in the biological sample; (10) a kit comprising a container  
 CC containing the novel polynucleotide, its degenerate variant or fragment,  
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the  
 CC genetically engineered host cell under conditions suitable to produce the  
 CC polypeptide from the culture. (1) can be used in gene therapy. The

CC polynucleotides, polypeptides, antibodies and compositions of the present  
 CC invention can be used for treating and diagnosing diseases, drug  
 CC screening assays and monitoring of effects during drug clinical trials.  
 CC The polynucleotides are useful for expressing and detecting Alloiooccus  
 CC otitis. The present sequence represents an Alloiooccus otitis  
 CC antigen protein from the present invention.  
 XX  
 SQ Sequence 264 AA;  
 Query Match 65.6%; Score 40; DB 6; Length 264;  
 Best Local Similarity 72.7%; Pred. NO. 72;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 QEKVLQKLGA 13  
 | | | | |  
 Db 117 QEKVAQLGKS 127

RESULT 38  
 ADB07516  
 ID ADB07516 standard; protein; 297 AA.  
 AC ADB07516;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Alloiooccus otitis antigenic protein SEQ ID NO:1456.  
 XX  
 KW Alloiooccus otitis; antigenic protein; immunogenic; immunisation;  
 XX gene therapy; Gram-positive bacterium; infection.  
 XX  
 OS Alloiooccus otitis.  
 XX  
 FN WO2003048304-A2.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PF 25-NOV-2002; 2002WO-US036123.  
 XX  
 PR 29-NOV-2001; 2001US-0333777P.  
 PR 18-NOV-2002; 2002US-0426742P.  
 XX  
 PA (AMHP) WYETH HOLDINGS CORP.  
 XX  
 PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;  
 XX  
 DR WPI; 2003-505284/47.  
 DR N-PSDB; ADB07515.  
 XX  
 PT New Alloiooccus otitis polynucleotides and polypeptides, useful for  
 PT treating and diagnosing diseases, drug screening assays and monitoring of  
 PT effects during drug clinical trials.  
 XX  
 PS Claim 33; SEQ ID NO 1456; 1019pp; English.  
 XX  
 CC The present invention describes an isolated polynucleotide (1) of  
 CC Alloiooccus otitis genomic DNA, which encodes an antigenic protein.  
 CC Alloiooccus otitis is a Gram-positive bacterium. Also described: (1)  
 CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an  
 CC expression vector comprising the novel isolated polynucleotide (1); its  
 CC complement, degenerate variant or fragment; (3) a genetically engineered  
 CC host cell, transfected, transformed or infected with the vector of (2);  
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
 CC composition comprising the polypeptide, its complement, biological  
 CC equivalent or fragment, or the polynucleotide that is comprised in the  
 CC expression vector; (6) a pharmaceutical composition comprising the  
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
 CC of the polypeptides of (1); their biological equivalent or fragment; (8)  
 CC immunising against Alloiooccus otitis by administering to a host the  
 CC immunogenic composition; (9) detecting and/or identifying Alloiooccus  
 CC otitis in the biological sample; (10) a kit comprising a container  
 CC containing the novel polynucleotide, its degenerate variant or fragment,  
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the

CC genetically engineered host cell under conditions suitable to produce the  
 CC polypeptide from the culture. (I) can be used in gene therapy. The  
 CC polynucleotides, polypeptides, antibodies and compositions of the present  
 CC invention can be used for treating and diagnosing diseases, drug  
 CC screening assays and monitoring of effects during drug clinical trials.  
 CC The polynucleotides are useful for expressing and detecting *Alloicoccus*  
 CC otitidis. The present sequence represents an *Alloicoccus* otitidis  
 CC antigen protein from the present invention.  
 XX  
 SQ Sequence 297 AA;

Query Match 65.6%; Score 40; DB 6; Length 297;  
 Best Local Similarity 72.7%; Pred. No. 82;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 QEKVLQKLGA 13  
 |||||:|:  
 DB 150 QEKVAQELGKS 160

RESULT 39

AAE15583  
 ID AAE15583 standard; protein; 154 AA.

XX AC AAE15583;

XX DT 12-MAR-2002 (first entry)

XX DE Mutated Glycine max ankyrin-related protein 2 (GMAdelta4).

XX KW Glycine max ankyrin-related protein; GMA2; GmsUT1; sugar transport;  
 XX KW sucrose/proton symporter; soybean ankyrin-related protein; SAR; ANK;  
 XX KW sugar allocation; nutritional value; mutant; mutein; GMAdelta4.

XX OS Glycine max.  
 XX OS Synthetic.

XX PN WO200188139-A2.

XX PD 22-NOV-2001.

XX PF 11-MAY-2001; 2001WO-US015315.

XX PR 12-MAY-2000; 2000US-0203974P.

XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX PI Grimes HD, Elmer AM, Murphy KA;

XX DR WPI; 2002-062385/08.

XX PT New purified protein having Glycine max ankyrin-related (GMA) protein  
 XX PT biological activity, useful to alter GMA levels in plants to confer  
 XX PT altered sugar transport and/or altered sugar allocation properties.

XX PS Example 2; Pg -; 60pp; English.

XX CC The invention relates to (soybean) Glycine max (Gm) ankyrin (ANK) -  
 XX CC related proteins 1 and 2 referred to as GMA1 and GMA2 and nucleic acid  
 XX CC molecules encoding them. GMA proteins also known as soybean ankyrin-  
 XX CC related (SAR) proteins interact with Gm sucrose/H+ (proton) symporter  
 XX CC designated as (SUT1). Manipulating the expression of GmsUT1 and GMA in  
 XX CC plants is useful to confer altered sugar transport and/or altered sugar  
 XX CC allocation properties. Alteration of GMA protein levels in plants could  
 XX CC be used to increase the nutritional value of plant tissues, for instance  
 XX CC plant seeds or grain. The present sequence is GMA2 mutant protein. This  
 XX CC sequence is obtained from the wild type GMA2 protein by deleting all the  
 XX CC four carboxy terminal ANK repeats. Note: This sequence is not found in  
 XX CC the specification but is derived from the wild type GMA2 protein shown as  
 XX CC SEQ ID NO:6 (AAE15309) in Fig 2 of the specification

XX SQ Sequence 154 AA;

Query Match 63.9%; Score 39; DB 5; Length 154;  
 Best Local Similarity 80.0%; Pred. No. 62;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 EKVLOKLGA 13  
 |||||:|:  
 DB 106 EKVLOKLGA 115

RESULT 40

AAE15580

ID AAE15580 standard; protein; 179 AA.

XX AC AAE15580;

XX DT 12-MAR-2002 (first entry)

XX DE Mutated Glycine max ankyrin-related protein 1 (GMAdelta4).

XX KW Glycine max ankyrin-related protein; GMA1; GmsUT1; sugar transport;  
 XX KW sucrose/proton symporter; soybean ankyrin-related protein; SAR; ANK;  
 XX KW sugar allocation; nutritional value; mutant; mutein; GMAdelta4.

XX OS Glycine max.

XX OS Synthetic.

XX PN WO200188139-A2.

XX PD 22-NOV-2001.

XX PF 11-MAY-2001; 2001WO-US015315.

XX PR 12-MAY-2000; 2000US-0203974P.

XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX PI Grimes HD, Elmer AM, Murphy KA;

XX DR WPI; 2002-062385/08.

XX PT New purified protein having Glycine max ankyrin-related (GMA) protein  
 XX PT biological activity, useful to alter GMA levels in plants to confer  
 XX PT altered sugar transport and/or altered sugar allocation properties.

XX PS Example 2; Pg -; 60pp; English.

XX CC The invention relates to (soybean) Glycine max (Gm) ankyrin (ANK) -  
 XX CC related proteins 1 and 2 referred to as GMA1 and GMA2 and nucleic acid  
 XX CC molecules encoding them. GMA proteins also known as soybean ankyrin-  
 XX CC related (SAR) proteins interact with Gm sucrose/H+ (proton) symporter  
 XX CC designated as (SUT1). Manipulating the expression of GmsUT1 and GMA in  
 XX CC plants is useful to confer altered sugar transport and/or altered sugar  
 XX CC allocation properties. Alteration of GMA protein levels in plants could  
 XX CC be used to increase the nutritional value of plant tissues, for instance  
 XX CC plant seeds or grain. The present sequence is GMA1 mutant protein. This  
 XX CC sequence is obtained from the wild type GMA1 protein by deleting all the  
 XX CC four carboxy terminal ANK repeats. Note: This sequence is not found in  
 XX CC the specification but is derived from the wild type GMA1 protein shown as  
 XX CC SEQ ID NO:4 (AAE15308) in Fig 2 of the specification

XX SQ Sequence 179 AA;

Query Match 63.9%; Score 39; DB 5; Length 179;  
 Best Local Similarity 80.0%; Pred. No. 73;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 EKVLOKLGA 13  
 |||||:|:  
 DB 131 EKVLOKLGA 140

RESULT 41

AAE15579



XW Glycine max ankyrin-related protein; GMA1; GmsUT1; sugar transport;  
 KW sucrose/proton symporter; soybean ankyrin-related protein; SAR; ANK;  
 XX sugar allocation; nutritional value.  
 OS Glycine max.  
 FH Key Location/Qualifiers  
 FT Misc-difference 3  
 FT /note= "Encoded by GGT; this residue is given as Gly in  
 FT the sequence shown as SEQ ID NO:4 in page 55-56 of the  
 FT specification"  
 FT 170..196  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif #4"  
 FT 201..230  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif"  
 FT 234..263  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif"  
 FT 267..293  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif"  
 FT 292..293  
 FT Misc-difference 293  
 FT /note= "Encoded by CTG"  
 FT Misc-difference 293  
 FT /note= "This residue is absent in the sequence shown as  
 FT SEQ ID NO:4 in page 55-56 of the specification"  
 FT 293  
 PN WO200188139-A2.  
 XX 22-NOV-2001.  
 PD 22-NOV-2001.  
 XX 11-MAY-2001; 2001WO-US015315.  
 XX 12-MAY-2000; 2000US-0203974P.  
 PR (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA Grimes HD, Elmer AM, Murphy KA;  
 PI WPI; 2002-062385/08.  
 DR N-PSDB; AAD24774.  
 XX New purified protein having Glycine max ankyrin-related (GMA) protein  
 PT biological activity, useful to alter GMA levels in plants to confer  
 PT altered sugar transport and/or altered sugar allocation properties.  
 PS Claim 1; Fig 2; 60pp; English.  
 XX The invention relates to (soybean) Glycine max (Gm) ankyrin (ANK)-  
 CC related proteins 1 and 2 referred to as GMA1 and GMA2 and nucleic acid  
 CC molecules encoding them. GMA proteins also known as soybean ankyrin-  
 CC related (SAR) proteins interact with Gm sucrose/H+ (proton) symporter  
 CC designated as (SUT1). Manipulating the expression of GmsUT1 and GMA in  
 CC plants is useful to confer altered sugar transport and/or altered sugar  
 CC allocation properties. Alteration of GMA protein levels in plants could  
 CC be used to increase the nutritional value of plant tissues, for instance  
 CC plant seeds or grain. The present sequence is GMA1 protein  
 XX Sequence 293 AA;  
 SQ  
 Query Match 63.9%; Score 39; DB 5; Length 293;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 EKVLOKLGKA 13  
 Db 131 EEVLQKLGQA 140  
 RESULT 44  
 AAE15582  
 ID AAE15582 standard; protein; 295 AA.  
 XX AAE15582;  
 AC AAE15582;  
 XX

DT 12-MAR-2002 (first entry)  
 XX Alternative version of Glycine max ankyrin-related protein 2 (GMA2).  
 DE Glycine max ankyrin-related protein; GMA2; GmsUT1; sugar transport;  
 XX sucrose/proton symporter; soybean ankyrin-related protein; SAR; ANK;  
 KW sugar allocation; nutritional value.  
 XX Glycine max.  
 OS  
 FH Key Location/Qualifiers  
 FT Domain 145..171  
 FT /note= "Ankyrin (ANK) repeat motif #4"  
 FT 176..205  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif"  
 FT 209..238  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif"  
 FT 242..268  
 FT Domain  
 FT /note= "Ankyrin (ANK) repeat motif"  
 FT 267  
 FT Misc-difference 267  
 FT /note= "Encoded by CTG TAG"  
 FT 268..295  
 FT Misc-difference 268  
 FT /note= "These residues are absent in the sequence shown  
 FT as SEQ ID NO:6 in Fig 2 of the specification"  
 FT 268  
 FT Misc-difference 268  
 FT /note= "Encoded by TTG TAA"  
 FT 288  
 FT Misc-difference 288  
 FT /note= "Encoded by TGG TGA"  
 FT 288  
 PN WO200188139-A2.  
 XX 22-NOV-2001.  
 PD 22-NOV-2001.  
 XX 11-MAY-2001; 2001WO-US015315.  
 XX 12-MAY-2000; 2000US-0203974P.  
 PR (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PA Grimes HD, Elmer AM, Murphy KA;  
 PI WPI; 2002-062385/08.  
 DR N-PSDB; AAD24775.  
 XX New purified protein having Glycine max ankyrin-related (GMA) protein  
 PT biological activity, useful to alter GMA levels in plants to confer  
 PT altered sugar transport and/or altered sugar allocation properties.  
 PS Claim 1; Page 56-57; 60pp; English.  
 XX The invention relates to (soybean) Glycine max (Gm) ankyrin (ANK)-  
 CC related proteins 1 and 2 referred to as GMA1 and GMA2 and nucleic acid  
 CC molecules encoding them. GMA proteins also known as soybean ankyrin-  
 CC related (SAR) proteins interact with Gm sucrose/H+ (proton) symporter  
 CC designated as (SUT1). Manipulating the expression of GmsUT1 and GMA in  
 CC plants is useful to confer altered sugar transport and/or altered sugar  
 CC allocation properties. Alteration of GMA protein levels in plants could  
 CC be used to increase the nutritional value of plant tissues, for instance  
 CC plant seeds or grain. The present sequence is GMA2 protein. Note: This  
 CC sequence is stated to be the same as that shown as SEQ ID NO:6 (AAE15309)  
 CC in Fig 2 of the specification. However the sequences differ at several  
 CC positions  
 XX Sequence 295 AA;  
 SQ  
 Query Match 63.9%; Score 39; DB 5; Length 295;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 EKVLOKLGKA 13  
 Db 106 EEVLQKLGQA 115



## RESULT 45

AA081620  
ID AA081620 standard; protein; 506 AA.

XX AC  
XX AA081620;  
XX DT 24-MAY-2000 (first entry)

XX DE Streptococcus pneumoniae type 4 protein sequence #120.

XX KW Streptococcus pneumoniae; vaccine; screening; protein antigen;  
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;  
KW pneumococcal disease.

XX OS Streptococcus pneumoniae.

XX PN WO200006737-A2.

XX PD 10-FEB-2000.

XX PF 27-JUL-1999; 99WO-GB002451.

XX PR 27-JUL-1998; 98GB-00016337.

XX PR 19-MAR-1999; 99US-0125164P.

XX PA (MICR-) MICROBIAL TECHNIQS LTD.

XX PI Gilbert CFG, Hansbro PM;

XX DR WPI; 2000-195300/17.

XX PT New Streptococcal protein, useful as a vaccine, for diagnosis of  
PT pneumococcal diseases and for screening agents capable of antagonizing or  
PT inhibiting expression of the protein.  
XX PS Claim 1; Page 84; 108pp; English.

XX CC AA081501 to AA081679 represent specifically claimed protein sequences  
CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent  
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
CC The sequences have antibacterial and antiinflammatory properties. The  
CC protein sequences, and fragments of them, are useful as immunogens and/or  
CC antigens. The nucleotide sequences can be used in vaccines and in  
CC diagnostic assays. The proteins and nucleotides can be useful for the  
CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
CC useful for screening an agent capable of antagonising, inhibiting or  
CC interfering with the function or expression of the proteins in which the  
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection  
CC and meningitis. AA05591 to AA05614 represent primers used in the  
CC exemplification of the present invention

XX SQ Sequence 506 AA;

Query Match 63.9%; Score 39; DB 3; Length 506;

Best Local Similarity 88.9%; Pred. No. 2.1e+02;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 KVLQKLGKA 13  
|||:|

DB 3 KVLQKVGKA 11  
|||:|

## RESULT 46

AAU38014

ID AAU38014 standard; protein; 510 AA.

XX AC AAU38014;

XX DT 14-FEB-2002 (first entry)

XX DE Streptococcus pneumoniae cellular proliferation protein #443.

KW

KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

XX antibacterial; drug design.

OS Streptococcus pneumoniae.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009180.

XX PR 21-MAR-2000; 2000US-0191078P.

XX PR 23-MAY-2000; 2000US-0206848P.

XX PR 26-MAY-2000; 2000US-0207727P.

XX PR 23-OCT-2000; 2000US-0242578P.

XX PR 27-NOV-2000; 2000US-0253625P.

XX PR 22-DEC-2000; 2000US-0257931P.

XX PR 16-FEB-2001; 2001US-0269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX DR WPI; 2001-611495/70.

XX DR N-PSDB; AAS55873.

XX PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.

XX PS Example 3; SEQ ID NO 13607; 511pp; English.

XX CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 510 AA;

Query Match 63.9%; Score 39; DB 4; Length 510;

Best Local Similarity 88.9%; Pred. No. 2.1e+02;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 KVLQKLGKA 13  
|||:|

DB 8 KVLQKVGKA 16  
|||:|

## RESULT 47

AAU37881

ID AAU37881 standard; protein; 510 AA.

XX AC AAU37881;

XX DT 14-FEB-2002 (first entry)

XX DE Streptococcus pneumoniae cellular proliferation protein #310.

XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

KW antibacterial; drug design.  
 XX Streptococcus pneumoniae.  
 OS WO200170955-A2.  
 PN 27-SEP-2001.  
 XX  
 PD  
 XX  
 FF 21-MAR-2001; 2001WO-US009180.  
 XX  
 PR 21-MAR-2000; 2000US-0191078P.  
 PR 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207727P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;  
 PI Yamamoto RT, Xu HH;  
 XX WPI; 2001-611495/70.  
 DR N-PSDB; AAS55740.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 XX  
 PS Example 3; SEQ ID NO 13474; 511pp; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 510 AA;  
 Query Match 63.9%; Score 39; DB 4; Length 510;  
 Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 KVLQKLGKA 13  
 |||||:  
 Db 8 KVLQKVGKA 16  
 RESULT 48  
 ABU02157  
 ID ABU02157 standard; protein; 510 AA.  
 XX  
 AC ABU02157;  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 11-FEB-2003 (first entry)  
 XX  
 DE S. pneumoniae type 4 strain protein from coding region #1734.  
 XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 KW

KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 XX gene therapy; vaccine.  
 OS Streptococcus pneumoniae; type 4 strain.  
 PN WO200277021-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 FF 27-MAR-2002; 2002WO-IB002163.  
 XX  
 PR 27-MAR-2001; 2001GB-00007658.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Masignani V, Tettelin H, Fraser C;  
 XX WPI; 2003-040579/03.  
 DR N-PSDB; ABX07446.  
 XX  
 PT New proteins and nucleic acid molecules from *Streptococcus pneumoniae*,  
 PT useful as medicaments for treating or preventing a disease or infection  
 PT due to *Streptococcus bacteria*, such as pneumonia, sepsis, otitis media or  
 PT ear infection.  
 XX  
 PS Claim 1; SEQ ID NO 3468; 56pp; English.  
 XX  
 CC The invention relates to a protein comprising or having at least 50%  
 CC identity to any of the 2469 amino acid sequences, identified in the  
 CC specification (available on a computer readable format), or its fragment,  
 CC expressed from 2469 of 2489 identified DNA coding regions from the  
 CC *Streptococcus pneumoniae* type 4 strain genomic sequence appearing as  
 CC AB356454. Also included are an antibody which binds one of the proteins,  
 CC treating a patient by administering the protein, DNA or antibody (in a  
 CC composition), a kit comprising first and second primers, which are the  
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a  
 CC sequence not defined in the specification, for amplifying a target  
 CC and the first primer is substantially complementary to the target sequence  
 CC and the second primer is substantially complementary to the target sequence  
 CC the target sequence, and where the parts of the primers having  
 CC substantial complementarity define the termini of the target sequence to  
 CC be amplified, assay comprising contacting a test compound with the  
 CC protein, and determining whether the test compound binds to the protein  
 CC and a *Streptococcus pneumoniae* bacterium, where one or more genes  
 CC encoding the proteins has been rendered inactive. The proteins, nucleic  
 CC acid molecules, antibody and compositions are useful as medicaments for  
 CC treating or preventing a disease or infection due to *Streptococcus*  
 CC bacteria, particularly *S. pneumoniae*, such as pneumonia, sepsis, otitis  
 CC media or ear infection. They are also useful in developing vaccines,  
 CC diagnostics and antibiotics. The methods are useful for identifying  
 CC immunodominant proteins. The present sequence is one of the 2469 proteins  
 CC expressed by the identified coding regions from the genomic sequence.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 510 AA;  
 Query Match 63.9%; Score 39; DB 6; Length 510;  
 Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 KVLQKLGKA 13  
 |||||:  
 Db 8 KVLQKVGKA 16  
 RESULT 49  
 ABU46197  
 ID ABU46197 standard; protein; 510 AA.

XX AC AEU46197;  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #31724.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Streptococcus pneumoniae.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu RH;  
XX DR WPI: 2003-029926/02.  
XX DR N-PSDB; ACA50067.  
XX XX  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX XX  
XX PS Claim 25; SEQ ID NO 74121; 1766pp; English.  
XX XX  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 510 AA;

Query Match 63.9%; Score 39; DB 6; Length 510;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;

XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #31724.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Streptococcus pneumoniae.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu RH;  
XX DR WPI: 2003-029926/02.  
XX DR N-PSDB; ACA50067.  
XX XX  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX XX  
XX PS Claim 25; SEQ ID NO 74121; 1766pp; English.  
XX XX  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 510 AA;

Query Match 62.3%; Score 38; DB 4; Length 80;  
Best Local Similarity 63.6%; Pred. No. 47;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KVLQKLGKA 13  
DB 8 KVLQKVGKA 16  
RESULT 50  
AAU58951  
ID AAU58951 standard; protein; 80 AA.  
XX AC AAU58951;  
XX DT 13-FEB-2002 (first entry)  
XX DE Propionibacterium acnes immunogenic protein #19847.  
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX KW dermatological; osteopathic; neuroprotectant.  
XX OS Propionibacterium acnes.  
XX PN WO200181581-A2.  
XX PD 01-NOV-2001.  
XX PF 20-APR-2001; 2001WO-US012865.  
XX PR 21-APR-2000; 2000US-0199047P.  
XX PR 02-JUN-2000; 2000US-0208841P.  
XX PR 07-JUL-2000; 2000US-0216747P.  
XX PA (CORI-) CORIXA CORP.  
XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX DR WPI: 2001-616774/71.  
XX DR N-PSDB; AAS59596.  
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for  
XX PT vaccinating against and diagnosing infections, especially useful for  
XX PT treating acne vulgaris.  
XX PS Example 1; SEQ ID NO 20146; 1069pp; English.  
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
XX CC polypeptides. The proteins and their associated DNA sequences are used in  
XX CC the treatment, prevention and diagnosis of medical conditions caused by  
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
XX CC P. acnes is also involved in infections of bone, joints and the central  
XX CC nervous system, however it is particularly involved in the inflammatory  
XX CC lesions associated with acne vulgaris. A method for detecting the  
XX CC presence or absence of P. acnes in a patient comprises contacting a  
XX CC sample with a binding agent that binds to the proteins of the invention  
XX CC and determining the amount of bound protein in the sample. The  
XX CC polypeptides may be used as antigens in the production of antibodies  
XX CC specific for P. acnes proteins. These antibodies can be used to  
XX CC downregulate expression and activity of P. acnes polypeptides and  
XX CC therefore treat P. acnes infections. The antibodies may also be used as  
XX CC diagnostic agents for determining P. acnes presence, for example, by  
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
XX CC this patent did not form part of the printed specification, but was  
XX CC obtained in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 80 AA;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 11  
 Db 66 RQEEVMORLG 76

Search completed: March 4, 2004, 17:44:17  
 Job time : 47.3548 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:41:21 ; Search time 13.4194 Seconds  
(without alignments)  
50.013 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_23\_35  
Perfect score: 61  
Sequence: 1 RAEKVKLQKGA 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents AA:\*

1: /cgm2\_6/ptodata/2/iaa/5A.COMB.pep:\*

2: /cgm2\_6/ptodata/2/iaa/5B.COMB.pep:\*

3: /cgm2\_6/ptodata/2/iaa/5A.COMB.pep:\*

4: /cgm2\_6/ptodata/2/iaa/5B.COMB.pep:\*

5: /cgm2\_6/ptodata/2/iaa/PTCUS.COMB.pep:\*

6: /cgm2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	61	100.0	434	US-08-630-915A-22
2	61	100.0	451	US-08-435-454-4
3	61	100.0	451	US-08-652-972A-4
4	61	100.0	451	US-08-919-145-6
5	61	100.0	451	US-08-870-126-4
6	61	100.0	451	US-09-344-889-6
7	61	100.0	451	US-09-445-247-4
8	61	100.0	451	PCT-US96-06231A-4
9	40	85.6	502	US-09-252-991A-22157
10	38	62.3	128	US-09-134-000C-3570
11	38	62.3	427	US-08-065-844A-2
12	38	62.3	487	US-09-543-681A-5185
13	38	62.3	490	US-09-328-352-4282
14	38	62.3	501	US-09-540-236-2078
15	38	62.3	562	US-09-489-039A-12840
16	37	60.7	209	US-09-370-838-182
17	37	60.7	447	US-09-205-258-408
18	36	59.0	279	US-09-134-000C-4472
19	36	59.0	285	US-08-410-540-2
20	36	59.0	285	US-08-659-254-18
21	36	59.0	323	US-08-522-991A-21134
22	36	59.0	410	US-09-880-137-3
23	36	59.0	418	US-09-880-137-1
24	36	59.0	418	US-09-880-137-2
25	35	57.4	153	US-09-107-532A-6174
26	35	57.4	228	US-09-702-953B-10
27	35	57.4	382	US-09-880-137-8

399	57.4	35	4	US-09-107-532A-6398	Sequence 6398, Ap
409	57.4	35	4	US-09-880-137-4	Sequence 4, Appli
409	57.4	35	4	US-09-880-137-7	Sequence 7, Appli
468	57.4	35	4	US-09-328-352-6321	Sequence 6321, Ap
574	57.4	35	4	US-09-702-953B-4	Sequence 4, Appli
627	57.4	35	4	US-09-702-953B-3	Sequence 3, Appli
647	57.4	35	4	US-09-702-953B-2	Sequence 2, Appli
227	56.6	34	4	US-09-914-259-61	Sequence 61, Appl
215	55.7	34	4	US-09-107-532A-3730	Sequence 3730, Ap
589	55.7	34	4	US-09-107-532A-6086	Sequence 6086, Ap
80	54.1	33	4	US-08-808-277A-3	Sequence 3, Appli
80	54.1	33	4	US-08-746-160-3	Sequence 3, Appli
276	54.1	33	4	US-08-659-254-2	Sequence 2, Appli
284	54.1	33	4	US-08-538-360-2	Sequence 2, Appli
296	54.1	33	4	US-08-134-001C-4230	Sequence 4230, Ap
423	54.1	33	4	US-08-855-910-13	Sequence 13, Appl
427	54.1	33	4	US-09-134-000C-5142	Sequence 5142, Ap
54.1	54.1	33	4	US-08-477-451-18	Sequence 18, Appl
496	54.1	33	4	US-08-313-075A-50	Sequence 50, Appl
586	54.1	33	4	US-09-519-232-8	Sequence 8, Appli
822	54.1	33	4	US-09-107-532A-6359	Sequence 6359, Ap
826	54.1	33	4	US-09-543-681A-5053	Sequence 5053, Ap
1105	54.1	33	4	US-09-540-236-3299	Sequence 3299, Ap
749	53.3	32	4	US-09-562-737-100	Sequence 100, App
90	52.5	32	4	US-09-621-976-5404	Sequence 5404, Ap
96	52.5	32	4	US-09-636-215-827	Sequence 827, App
96	52.5	32	4	US-09-685-166A-827	Sequence 827, App
100	52.5	32	4	US-09-198-452A-1076	Sequence 1076, Ap
165	52.5	32	4	US-09-896-866B-13	Sequence 13, Appl
165	52.5	32	4	US-09-855-341-4	Sequence 4, Appli
166	52.5	32	4	US-09-134-000C-4121	Sequence 4121, Ap
170	52.5	32	4	US-09-855-341-16	Sequence 16, Appl
171	52.5	32	4	US-09-328-352-8227	Sequence 8227, Ap
224	52.5	32	4	US-09-636-215-825	Sequence 825, App
224	52.5	32	4	US-09-685-166A-825	Sequence 825, App
227	52.5	32	4	US-09-896-866B-17	Sequence 17, Appl
227	52.5	32	4	US-09-855-341-8	Sequence 8, Appli
310	52.5	32	4	US-08-363-255-6	Sequence 6, Appli
310	52.5	32	4	US-09-047-026A-25	Sequence 25, Appl
319	52.5	32	4	US-09-107-532A-4723	Sequence 4723, Ap
332	52.5	32	4	US-09-252-991A-30897	Sequence 30897, A
344	52.5	32	4	US-09-415-277C-5	Sequence 5, Appli
353	52.5	32	4	US-09-543-681A-5613	Sequence 5613, Ap
355	52.5	32	4	US-09-134-001C-3622	Sequence 3622, Ap
388	52.5	32	4	US-09-880-137-5	Sequence 5, Appli
388	52.5	32	4	US-09-880-137-6	Sequence 6, Appli
389	52.5	32	4	US-09-800-170-18	Sequence 18, Appl
399	52.5	32	4	US-09-252-991A-27261	Sequence 27261, A
427	52.5	32	4	US-09-107-532A-5530	Sequence 5530, Ap
442	52.5	32	4	US-08-363-255-4	Sequence 4, Appli
442	52.5	32	4	US-08-363-255-5	Sequence 5, Appli
445	52.5	32	4	US-08-363-255-12	Sequence 12, Appl
445	52.5	32	4	US-09-134-000C-4191	Sequence 4191, Ap
453	52.5	32	4	US-09-092-315-2	Sequence 2, Appli
486	52.5	32	4	US-09-733-524A-2	Sequence 2, Appli
486	52.5	32	4	US-09-543-681A-6439	Sequence 6439, Ap
508	52.5	32	4	US-09-457-040B-16	Sequence 16, Appl
517	52.5	32	4	US-09-134-000C-5318	Sequence 5318, Ap
577	52.5	32	4	US-07-820-154A-30	Sequence 30, Appl
577	52.5	32	4	US-08-663-566A-11	Sequence 11, Appl
577	52.5	32	4	US-08-097-554A-30	Sequence 30, Appl
577	52.5	32	4	US-08-023-610-11	Sequence 11, Appl
577	52.5	32	4	US-08-288-065A-11	Sequence 11, Appl
577	52.5	32	4	US-08-362-240A-11	Sequence 11, Appl
577	52.5	32	4	US-08-480-640A-30	Sequence 30, Appl
577	52.5	32	4	US-08-295-802-30	Sequence 30, Appl
577	52.5	32	4	US-08-804-372A-9	Sequence 9, Appli
577	52.5	32	4	US-08-488-237A-30	Sequence 30, Appl
577	52.5	32	4	US-08-375-992A-30	Sequence 30, Appl
577	52.5	32	4	US-08-472-679H-30	Sequence 30, Appl
577	52.5	32	4	PCT-US93-00324-30	Sequence 30, Appl
577	52.5	32	4	PCT-US95-10245-11	Sequence 11, Appl

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York

; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-435-454-4

Query Match 100.0%; Score 61; DB 1; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 |||||  
 Db 21 RAQEKVLQKLGKA 33

## RESULT 3

US-08-652-972A-4  
 ; Sequence 4, Application US/08652972A  
 ; Patent No. 5723581

## GENERAL INFORMATION:

; APPLICANT: Prendergast, George C.  
 ; APPLICANT: Sakamuro, Daitoku  
 ; TITLE OF INVENTION: Murine and Human Box-Dependent  
 ; TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P O Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/652,972A  
 ; FILING DATE: 24-MAY-1996  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/435,454  
 ; FILING DATE: 05-MAY-1995

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bak, Mary E.  
 ; REGISTRATION NUMBER: 31,215  
 ; REFERENCE/DOCKET NUMBER: WST60BUSA

; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear

; MOLECULE TYPE: protein  
 US-08-652-972A-4

Query Match 100.0%; Score 61; DB 1; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 |||||  
 Db 21 RAQEKVLQKLGKA 33

## RESULT 4

US-08-919-145-6  
 ; Sequence 6, Application US/08919145  
 ; Patent No. 5958753  
 ; GENERAL INFORMATION:

; APPLICANT: Prendergast, George C.  
 ; TITLE OF INVENTION: Bau, A Bin1 Interacting Protein, and  
 ; TITLE OF INVENTION: Uses Therefor  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P.O. Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/919,145  
 ; FILING DATE:

; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/025,482

; FILING DATE: 29-AUG-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kodroff, Cathy A.

; REGISTRATION NUMBER: 33,980  
 ; REFERENCE/DOCKET NUMBER: WST73AUSA  
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-08-919-145-6

Query Match 100.0%; Score 61; DB 2; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 |||||  
 Db 21 RAQEKVLQKLGKA 33

## RESULT 5

US-08-870-126-4  
 ; Sequence 4, Application US/08870126  
 ; Patent No. 6048702

## GENERAL INFORMATION:

; APPLICANT: Prendergast, George C.  
 ; APPLICANT: Sakamuro, Daitoku

; TITLE OF INVENTION: Murine and Human Box-Dependent  
 ; TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P O Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/870,126  
 ; FILING DATE:  
 ; CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,454  
FILING DATE: 05-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/652,972  
FILING DATE: 24-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kodroff, Cathy A.  
REGISTRATION NUMBER: 33,980  
REFERENCE/DOCKET NUMBER: WST60CUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-870-126-4

Query Match 100.0%; Score 61; DB 3; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLKGA 13  
|||||  
Db 21 RAQEKVLQKLKGA 33

RESULT 6  
US-09-344-889-6  
Sequence 6; Application US/09344889  
Patent No. 6140465  
GENERAL INFORMATION:  
APPLICANT: Prendergast, George C.  
TITLE OF INVENTION: Bau, A Binl Interacting Protein, and  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson and Howson  
STREET: Spring House  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,889  
FILING DATE:  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/919,145  
FILING DATE:  
FILING DATE: US 60/025,482  
FILING DATE: 29-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kodroff, Cathy A.  
REGISTRATION NUMBER: 33,980  
REFERENCE/DOCKET NUMBER: WST73AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-344-889-6

Query Match 100.0%; Score 61; DB 3; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLKGA 13  
|||||  
Db 21 RAQEKVLQKLKGA 33

RESULT 7  
US-09-445-247-4  
Sequence 4; Application US/09445247  
Patent No. 6410238  
GENERAL INFORMATION:  
APPLICANT: Wistar Institute of Anatomy & Biology  
Sakamuro, Daitoku  
PRENDERGAST, George C.

TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein  
(Binl) Compositions and Uses Therefor  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/445,247  
FILING DATE: 03-Dec-1999  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/870,126  
FILING DATE: 06-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60DPCT

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-445-247-4

Query Match 100.0%; Score 61; DB 4; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLKGA 13  
|||||  
Db 21 RAQEKVLQKLKGA 33

RESULT 8  
PCT-US96-06231A-4  
Sequence 4; Application PC/TUS9606231A  
GENERAL INFORMATION:  
APPLICANT: Wistar Institute of Anatomy & Biology  
TITLE OF INVENTION: Murine and Human Box-Dependent  
TITLE OF INVENTION: Myc-Interacting Protein (Binl) and Uses Therefor



NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06231A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,454  
FILING DATE: 05-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Maty E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60APCT  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-06231A-4

Query Match 100.0%; Score 61; DB 5; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 13  
DB 21 RAQKVLQKLGK 33

RESULT 9  
US-09-252-991A-22157  
Sequence 22157, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 22157  
LENGTH: 502  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22157

Query Match 65.6%; Score 40; DB 4; Length 502;  
Best Local Similarity 72.7%; Pred. No. 49;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQKVLQKLG 11  
DB 110 RAQKVLQKLG 120

RESULT 10  
US-09-134-000C-3570  
Sequence 3570, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032798-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 3570  
LENGTH: 128  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-3570

Query Match 62.3%; Score 38; DB 4; Length 128;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGK 12  
DB 2 EERILQKVGK 11

RESULT 11  
US-08-065-844A-2  
Sequence 2, Application US/08065844A  
Patent No. 6333168  
GENERAL INFORMATION:  
APPLICANT: Jessell, Thomas M.  
APPLICANT: Basler, Konrad  
APPLICANT: Yamada, Toshiya  
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF  
TITLE OF INVENTION: DORSALIN-1  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/065,844A  
FILING DATE: 19930520  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0576/40314  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOP U  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 427 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-065-844A-2

Query Match 62.3%; Score 38; DB 4; Length 427;  
Best Local Similarity 80.0%; Pred. No. 91;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3 QAEKVLQKLGK 12  
Db 282 QESVNLKLGK 291

RESULT 12

US-09-543-681A-5185  
; Sequence 5185, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 5185  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-5185

Query Match 62.3%; Score 38; DB 4; Length 487;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RAEKVLQKLGK 12  
Db 67 KAEKVFHQLGR 78

RESULT 13

US-09-328-352-4282  
; Sequence 4282, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4282  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4282

Query Match 62.3%; Score 38; DB 4; Length 490;  
Best Local Similarity 58.3%; Pred. No. 1e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RAEKVLQKLGK 12  
Db 89 KAEKVFSELGR 100

RESULT 14

US-09-540-236-2078  
; Sequence 2078, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; NUMBER OF SEQ ID NOS: 2000-04-04  
; SEQ ID NO 2078  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: M. catarrhalis  
US-09-540-236-2078

Query Match 62.3%; Score 38; DB 4; Length 501;  
Best Local Similarity 58.3%; Pred. No. 1.1e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RAEKVLQKLGK 12  
Db 99 KAEKVFSELGR 110

RESULT 15

US-09-489-039A-12840  
; Sequence 12840, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12840  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12840

Query Match 62.3%; Score 38; DB 4; Length 562;  
Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RAEKVLQKLGK 12  
Db 142 KAEKVFHQLGR 153

RESULT 16

US-09-370-838-182  
; Sequence 182, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Monamath, Roadon  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; FILE REFERENCE: 210121.475C1  
; CURRENT APPLICATION NUMBER: US/09/370,838  
; PRIOR FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/285,323  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 182  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-370-838-182

Query Match 60.7%; Score 37; DB 4; Length 209;  
 Best Local Similarity 80.0%; Pred. No. 68;  
 Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 AQEKVLQKLG 11  
 :|||||  
 Db 141 AREKVLQTLG 150

RESULT 17  
 US-09-205-258-408  
 ; Sequence 408, Application US/09205258  
 ; Patent No. 6525174  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young et al.  
 ; TITLE OF INVENTION: 207 Human Secreted Proteins  
 ; FILE REFERENCE: P2007P1  
 ; CURRENT APPLICATION NUMBER: US/09/205,258  
 ; EARLIER FILING DATE: 1998-12-04  
 ; EARLIER FILING DATE: 1998-06-04  
 ; EARLIER APPLICATION NUMBER: 60/048,885  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,375  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,881  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,880  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,896  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,020  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,876  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,895  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,884  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,894  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,971  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,964  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,882  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,899  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,893  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,900  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,901  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,892  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,915  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,019  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,970  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,972  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,916  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,373  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,875  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,374

; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,917  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,949  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,974  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,883  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,897  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,898  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,962  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,963  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,877  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,878  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/070,923  
 ; EARLIER FILING DATE: 1997-12-18  
 ; EARLIER APPLICATION NUMBER: 60/092,921  
 ; EARLIER FILING DATE: 1998-07-15  
 ; EARLIER APPLICATION NUMBER: 60/094,657  
 ; EARLIER FILING DATE: 1998-07-30  
 ; NUMBER OF SEQ ID NOS: 1227  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 408  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; NAME/KEY: SITE  
 ; LOCATION: (447)  
 ; OTHER INFORMATION: Xaa equals stop translation  
 US-09-205-258-408

Query Match 60.7%; Score 37; DB 4; Length 447;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11  
 :|||||  
 Db 378 AREKVLQTLG 387

RESULT 18  
 US-09-134-000C-4472  
 ; Sequence 4472, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4472  
 ; LENGTH: 279  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 US-09-134-000C-4472

Query Match 59.0%; Score 36; DB 4; Length 279;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 QKVLQKLG 11  
 |||:|:  
 Db 9 QKLIQELG 17

RESULT 19  
 US-08-410-540-2  
 ; Sequence 2, Application US/08410540  
 ; Patent No. 5807678  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Miller, Walter L.  
 ; APPLICANT: Lin, Dong  
 ; APPLICANT: Straus III, Jerome F.  
 ; TITLE OF INVENTION: IDENTIFICATION OF GENE MUTATIONS  
 ; TITLE OF INVENTION: ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
 ; STREET: 5 Palo Alto Square  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: US  
 ; ZIP: 94306-2155  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/410,540  
 ; FILING DATE: 23-MAR-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Neeley, Richard L.  
 ; REGISTRATION NUMBER: 30,092  
 ; REFERENCE/DOCKET NUMBER: UCAL-238/0005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415 853 5070  
 ; TELEFAX: 415 857 0663  
 ; TELEX: 380816CCOLEYPA  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 285 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-410-540-2

Query Match 59.0%; Score 36; DB 1; Length 285;  
 Best Local Similarity 87.5%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLG 12  
 |||:|:  
 Db 155 KVLQKIGK 162

RESULT 20  
 US-08-659-254-18  
 ; Sequence 18, Application US/08659254  
 ; Patent No. 6194555  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stocco, Douglas M.  
 ; APPLICANT: Clark, Dr. Barbara J.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION  
 ; TITLE OF INVENTION: OF STEROIDOGENESIS  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.  
 ; STREET: 1900 Frost Bank Plaza, 816 Congress Avenue  
 ; CITY: Austin  
 ; STATE: TX  
 ; COUNTRY: U.S.A.

ZIP: 78701  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/659,254  
 ; FILING DATE: 07-JUN-1996  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/538,960  
 ; FILING DATE: 04-OCT-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mayfield, Denise L.  
 ; REGISTRATION NUMBER: 33,732  
 ; REFERENCE/DOCKET NUMBER: 43375.0006  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 512/499-6200  
 ; TELEFAX: 512/499-6290  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 285 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-659-254-18

Query Match 59.0%; Score 36; DB 3; Length 285;  
 Best Local Similarity 87.5%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLG 12  
 |||:|:  
 Db 155 KVLQKIGK 162

RESULT 21  
 US-09-252-991A-21134  
 ; Sequence 21134, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 21134  
 ; LENGTH: 323  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-21134

Query Match 59.0%; Score 36; DB 4; Length 323;  
 Best Local Similarity 46.2%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGKA 13  
 |||:|:  
 Db 15 RSQDEVAERLQGA 27

RESULT 22  
 US-09-880-137-3  
 ; Sequence 3, Application US/09880137  
 ; Patent No. 6640025  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berstein, Gabriel

```

RESULT 6
Q9UKN4 PRELIMINARY; PRT; 564 AA.
AC Q9UKN4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Bridging integrator-2.
GN BIN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20366138; PubMed=10903846;
RA Ge X., Fundergast G.C.;
RT "Bin2, a Functionally Nonredundant Member of the BRAP1 Gene
RT Family.";
RL Genomics 67:210-220(2000).
DR EMBL; AF146531; AAD54227.1; -.
DR GO; GO:0006897; P:endocytosis; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR Pfam; PF03114; BAR_dom.
DR InterPro; IPR004148; BAR_1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 564 AA; 61747 MW; 3CB3791A56CE53BC CRC64;

Query Match 100.0%; Score 61; DB 4; Length 564;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
DB 23 RAQEKVLQKLGKA 35

RESULT 7
Q9UBW5 PRELIMINARY; PRT; 565 AA.
AC Q9UBW5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Breast cancer associated protein BRAP1.
GN BRAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX "Genomic structure and chromosome location of the BRAP1 gene.";
RA Miki Y., Saito H.;
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032710; BAA88125.1; -.
DR EMBL; AB032699; BAA88125.1; JOINED.
DR EMBL; AB032700; BAA88125.1; JOINED.
DR EMBL; AB032701; BAA88125.1; JOINED.
DR EMBL; AB032702; BAA88125.1; JOINED.
DR EMBL; AB032703; BAA88125.1; JOINED.
DR EMBL; AB032704; BAA88125.1; JOINED.
DR EMBL; AB032705; BAA88125.1; JOINED.
DR EMBL; AB032706; BAA88125.1; JOINED.
DR EMBL; AB032707; BAA88125.1; JOINED.
DR EMBL; AB032708; BAA88125.1; JOINED.
DR EMBL; AB032709; BAA88125.1; JOINED.
DR EMBL; AB032698; BAA88108.1; -.

```

```

DR Genew; HGNC:1053; BIN2.
DR GO; GO:0006897; P:endocytosis; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR Pfam; PF03114; BAR_1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 565 AA; 61902 MW; 28D1FB89A7779860 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
DB 23 RAQEKVLQKLGKA 35

RESULT 8
Q86VVO PRELIMINARY; PRT; 565 AA.
AC Q86VVO;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Breast cancer associated protein BRAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RC Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047686; AAH47686.1; -.
DR GO; GO:0006897; P:endocytosis; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR Pfam; PF03114; BAR_1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 565 AA; 61875 MW; B1D1FB89A7707EA3 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 565;

```

```

Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
DB 23 RAQEKVLQKLGKA 35

RESULT 9
Q7Q7F7
ID Q7Q7F7 PRELIMINARY; PRT; 686 AA.
AC Q7Q7F7;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=2238257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054718; AAH54718.1; -
KW Hypothetical protein.
SQ SEQUENCE 686 AA; 75013 MW; D292E24653A442A5 CRC64;

Query Match 100.0%; Score 61; DB 11; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
DB 19 RAQEKVLQKLGKA 31

RESULT 10
Q8N4G0
ID Q8N4G0 PRELIMINARY; PRT; 695 AA.
AC Q8N4G0;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Amphiphysin (Stiff-Man syndrome with breast cancer 128kD
DE autoantigen).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; BC034376; AAH34376.1; -
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR003017; Amphiphysin_1.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD003208; Amphiphysin_1; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 695 AA; 76229 MW; 4FC92E8B028BBFE CRC64;

Query Match 100.0%; Score 61; DB 4; Length 695;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
DB 19 RAQEKVLQKLGKA 31

RESULT 11
Q803I3
ID Q803I3 PRELIMINARY; PRT; 524 AA.
AC Q803I3;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to bridging integrator 2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044473; AAH44473.1; -
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 524 AA; 59280 MW; F4A2988732D153FF CRC64;

Query Match 93.4%; Score 57; DB 13; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12
DB 31 RAQEKVLQKLGK 42
```

```

RESULT 12
Q8WQ54 ID Q8WQ54 PRELIMINARY; PRT; 414 AA.
AC Q8WQ54; 2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Putative aggregation factor.
GN GBOCVAF.
OS Geodia cydonium (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Astrophorida; Geodidae; Geodia.
OX NCBI_TaxID=6047;
RN [1]
RP Schutze J., Kraak A., Diehl-Seifert B., Mueller W.E.G.;
RT "Cloning and Expression of the Putative Aggregation Factor from the
RT Marine Sponge Geodia cydonium."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ311598; CAC89970.1; -
DR GO; GO:0006897; P:cytoskeleton; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; PS50002; SH3; 1.
DR SH3 domain.
SQ SEQUENCE 414 AA; 46558 MW; 90CABDAB20152310 CRC64;

Query Match 80.3%; Score 49; DB 5; Length 414;
Best Local Similarity 69.2%; Pred. No. 2.8;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
Db 30 RVQEKMMQKLGKS 42

RESULT 13
Q9Y092 ID Q9Y092 PRELIMINARY; PRT; 602 AA.
AC Q9Y092;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Amphiphysin protein (ID19810P).
GN AMPH OR CG8604.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Bouchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

```

```

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Meleod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostreffi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20076404; PubMed=10607911;
RA Razaq A., Su Y., Mehren J.E., Mizuguchi K., Jackson A.P., Gay N.J.,
RA O'Kane C.J.;
RT "Characterization of the gene for Drosophila amphiphysin."
RL Gene 241:167-174(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX Zhang Y.Q., Xue H., Broadie K.S.;
RT "Dissecting the functions of amphiphysin."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AE003821; AAF58486.1; -
DR EMBL; AF171225; AAD46684.1; -
DR EMBL; AJ242855; CAB45188.1; -
DR EMBL; AY061278; AAL28826.1; -
DR HSP; P27986; 1PHT.
DR FlyBase; FBgn0027356; Amph.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0045202; C:synaptic junction; IDA.
DR GO; GO:0007269; P:neurotransmitter secretion; NAS.
DR GO; GO:0008104; P:protein localization; IMP.
DR GO; GO:0006937; P:regulation of muscle contraction; IMP.
DR GO; GO:0045313; P:rhabdomere membrane biogenesis; IMP.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; PS50002; SH3; 1.
DR SH3 domain.
SQ SEQUENCE 602 AA; 65901 MW; 86D1D4CBAE1B8F31 CRC64;

```

Query Match

77.0%; Score 47; DB 5; Length 602;

```

Best Local Similarity 75.0%; Pred. No. 8.9;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 12
DB 19 RAKEKILQNLGK 30

RESULT 14
Q9FE28 PRELIMINARY; PRT; 461 AA.
ID C9FE28;
AC Q9FE28;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0005A05.3 protein (P0436E04.22 protein).
GN P0005A05.3 OR P0436E04.22.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0005A05.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0436E04.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002863; BAB16899.1; -
DR EMBL; AP002818; BAB16339.1; -
DR Gramene; Q9FE28; -
SQ SEQUENCE 461 AA; 51281 MW; B5BCP4085F4EC37C CRC64;

Query Match 67.2%; Score 41; DB 10; Length 461;
Best Local Similarity 61.5%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 13
DB 145 RLKKAQLQKLGK 157

RESULT 15
P94544 PRELIMINARY; PRT; 570 AA.
ID P94544;
AC P94544;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein yshc.
GN YSHC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97217425; PubMed=9063446;
RA Goethel S.F.; Schmid R.; Wipat A.; Carter N.M.; Emerson P.T.;
RA Harwood C.R.; Marahiel M.A.;
RT "An internal PK506-binding domain is the catalytic core of the prolyl
isomerase activity associated with the Bacillus subtilis trigger
factor.";
RL Eur. J. Biochem. 244:59-65(1997).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F.; Ogasawara N.; Moser I.; Albertini A.M.; Alloni G.;
RA Azevedo V.; Bertero M.G.; Bessieres P.; Bolotin A.; Borchert S.;
RA Borriess R.; Boursier L.; Brans A.; Braun M.; Brignell S.C.; Bron S.;
RA Brouillet S.; Bruschi C.V.; Caldwell B.; Capuano V.; Carter N.M.;
RA Choi S.K.; Codani J.J.; Connerton I.F.; Cummings N.J.; Daniel R.A.;
RA Denizot F.; Devine K.M.; Dusterhoft A.; Ehrlich S.D.; Emerson P.T.;
RA Entian K.D.; Errington J.; Fabret C.; Ferrari E.; Foulger D.;
RA Fritz S.Y.; Glaser P.; Fujita Y.; Fuma S.; Gallizi A.; Galleron N.;
RA Guiseppi G.; Guy B.J.; Haga K.; Haiech J.; Harwood C.R.; Henaut A.;
RA Joris B.; Karamata D.; Kasahara Y.; Klaerr-Blanchard M.; Klein C.;
RA Kobayashi Y.; Koetter F.; Koningsstein G.; Krogh S.; Kumano M.;
RA Kurita K.; Lardius A.; Lardinois S.; Lauber J.; Lazarevic V.;
RA Lee S.M.; Levine A.; Liu H.; Masuda S.; Mauel C.; Medigue C.;
RA Medina N.; Mellado R.P.; Mizuno M.; Moestl D.; Nakai S.; Noback M.;
RA Noone D.; O'Reilly M.; Ogawa K.; Ogiwara A.; Oudega B.; Park S.H.;
RA Parro V.; Pohl T.M.; Portetelle D.; Porwollik S.; Prescott A.M.;
RA Presecan E.; Puig P.; Purnelle B.; Rapoport G.; Rey M.; Reynolds S.;
RA Rieger M.; Rivoita C.; Rocha E.; Roche B.; Rose M.; Sadale Y.;
RA Sato T.; Scanlan E.; Schleich S.; Schroeter R.; Scoffone F.;
RA Sekiguchi J.; Sekowska A.; Seror S.J.; Seror P.; Shin B.S.; Soldo B.;
RA Sorokin A.; Tacconi E.; Takagi T.; Takahashi H.; Takemaru K.;
RA Takeuchi M.; Tamakoshi A.; Tanaka T.; Terstra P.; Tognoni A.;
RA Toseato V.; Uchiyama S.; Vandenbol M.; Vannier F.; Vassarotti A.;
RA Viari A.; Wambutt R.; Wedler E.; Wedler H.; Weitzenecker T.;
RA Winters P.; Wipat A.; Yamamoto H.; Yamane K.; Yasumoto K.; Yata K.;
RA Yoshida K.; Yoshikawa H.F.; Zumstein E.; Yoshikawa H.; Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F.; Ogasawara N.; Yoshikawa H.; Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z75208; CAA99568.1; -
DR EMBL; Z99118; CABI4819.1; -
DR PIR; C69985; C69985.
DR HSSP; P06766; 13PB.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR002054; DNA_polX.
DR InterPro; IPR004013; PHP_C.
DR InterPro; IPR003141; PHP_N.
DR Pfam; PF02811; PHP_C; 1.
DR Pfam; PF02231; PHP_N; 1.
DR SMART; SM00481; POLIITAC; 1.
DR SMART; SM00483; POLXC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 570 AA; 64119 MW; 06AEA028BFA3D3F5 CRC64;

Query Match 67.2%; Score 41; DB 16; Length 570;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 13
DB 134 KSEKILQALGEA 146

RESULT 16
Q8E5M0 PRELIMINARY; PRT; 651 AA.
ID Q8E5M0;
AC Q8E5M0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN GBS0962.

```



OS Streptococcus agalactiae (serotype III).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=216495;  
 RN [1];

RP SEQUENCE FROM N.A.  
 RC STRAIN=NEM316 / Serotype III;  
 RX MEDLINE=22242508; PubMed=12354221;  
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,  
 RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,  
 RA Kunst F.;  
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
 RT invasive neonatal disease.";  
 RL Mol. Microbiol. 45:1459-1513 (2002).  
 DR EMBL; AL766848; CAD46621.1; -.  
 DR Sgallist; gbs0962; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003838; DUF214.  
 DR Pfam; PF02687; FtsX; 1.  
 KW Hypothetical protein, Complete proteome.  
 SQ SEQUENCE 651 AA; 74102 MW; 1383BB8A214335C5 CRC64;

Query Match 67.2%; Score 41; DB 16; Length 651;  
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13

Db 334 EKKLKLKLGKS 344

RESULT 17

Q8DX0 PRELIMINARY; PRT; 651 AA.  
 ID Q8DX0;  
 AC Q8DX0;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE ABC transporter, permease protein, putative.  
 GN SAG0975.

OS Streptococcus agalactiae (serotype V).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=216466;  
 RN [1];  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2603 V/R / Serotype V;  
 RX MEDLINE=2222988; PubMed=12200547;  
 RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,  
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,  
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,  
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,  
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,  
 RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarcelli M., Mora M.,  
 RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,  
 RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,  
 RA Fraser C.M.;  
 RT "Complete genome sequence and comparative genomic analysis of an  
 RT emerging human pathogen, serotype V Streptococcus agalactiae.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).  
 DR EMBL; AE014236; AAM99858.1; -.  
 DR TIGR; SAG0975; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR003838; DUF214.  
 DR Pfam; PF02687; FtsX; 1.  
 KW Complete proteome.

SQ SEQUENCE 651 AA; 74016 MW; A782A8A2143FDA8 CRC64;

Query Match 67.2%; Score 41; DB 16; Length 651;  
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13

Db 334 EKKLKLKLGKS 344

|||||:

RESULT 18

Q84PY5 PRELIMINARY; PRT; 874 AA.  
 ID Q84PY5;  
 AC Q84PY5;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE P0443G08.4 protein.  
 GN P0443G08.4.

OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1];

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC  
 RT clone:P0443G08.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP004461; BAC64982.1; -.  
 SQ SEQUENCE 874 AA; 97030 MW; C1AD2927BB2554D6 CRC64;

Query Match 67.2%; Score 41; DB 10; Length 874;  
 Best Local Similarity 61.5%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGKA 13

Db 145 RLKRLALQKMGKA 157

RESULT 19

Q943V8 PRELIMINARY; PRT; 876 AA.  
 ID Q943V8;  
 AC Q943V8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE B1146B04.1 protein (B1151H08.15 protein).  
 GN B1146B04.1 OR B1151H08.15.

OS Oryza sativa (Rice), and  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OX NCBI\_TaxID=4530, 39947;  
 RN [1];

RP SEQUENCE FROM N.A.

RC SPECIES=O. sativa; STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
 RT clone:B1146B04.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2];

RP SEQUENCE FROM N.A.

RC SPECIES=O. sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
 RT clone:B1151H08.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003205; BAB64572.1; -.  
 DR EMBL; AP003336; BAC00595.1; -.  
 DR Gramene; Q943V8; -.

SQ SEQUENCE 876 AA; 97403 MW; 4BFA6DF145ACC814 CRC64;

Query Match 67.2%; Score 41; DB 10; Length 876;  
 Best Local Similarity 61.5%; Pred. No. 1.4e+02;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
DB 145 RLKRALQKMGKA 157

RESULT 20

Q7XXG7 PRELIMINARY; PRT; 876 AA.  
ID Q7XXG7  
AC Q7XXG7  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE OSJNB0089K24.1 protein.  
GN OSJNB0089K24.1  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,  
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,  
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,  
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,  
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,  
RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,  
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,  
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,  
RA Zhang R.Q., Guan J.P., Hong G.P.  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL606509; CAD39392.1; --  
SQ SEQUENCE 876 AA; 97416 MW; 94D8EDF9DFE97475 CRC64;

Query Match 67.2%; Score 41; DB 10; Length 876;  
Best Local Similarity 61.5%; Pred. No. 1.4e+02;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
DB 145 RLKRALQKMGKA 157

RESULT 21

Q8SX64 PRELIMINARY; PRT; 1235 AA.  
ID Q8SX64  
AC Q8SX64  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-OCT-2002 (TREMELrel. 21, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE LD40094p.  
GN CG12734.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY094827; AAU1180.1; --  
DR FlyBase; FBgn0035411; CG12734.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR002017; Spectrin.  
DR PROSITE; PS00190; CYTOCHROME C; 1.  
SQ SEQUENCE 1235 AA; 142401 MW; C7C813ECBD86C58A CRC64;

Query Match 67.2%; Score 41; DB 5; Length 1235;  
Best Local Similarity 81.8%; Pred. No. 2e+02;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13  
DB 663 QSKELQKLGKA 673

RESULT 22

Q9VZT7 PRELIMINARY; PRT; 1381 AA.  
ID Q9VZT7  
AC Q9VZT7  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE CG12734 protein.  
GN CG12734.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gall R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Folsler C., Gabriellian A.E., Garg N.S., Galbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Maleshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy E., Murphy L., Nusskern D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Paillet J., Pan S., Pohlman J.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287;2195-2195(2000).  
DR EMBL; AS003477; AAF47730.1; --  
DR HSSP; P03069; IYSA.  
DR FlyBase; FBgn0035411; CG12734.

```

DR InterPro; IPR002017; Spectrin.
SQ SEQUENCE 1381 AA; 157512 MW; 2A75A71D74DD44E5 CRC64;

Query Match 67.2%; Score 41; DB 5; Length 1381;
Best Local Similarity 81.8%; Pred. No. 2.2e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13
Db 663 QSEKQLKLGKA 673

RESULT 23
Q9RWL3 PRELIMINARY; PRT; 176 AA.
AC Q9RWL3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Hypothetical protein DR0653.
GN DR0653.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCBI 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.; Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RT Science 286:1571-1577 (1999).
RL EMBL; AE001922; AAF10228.1; -.
DR PIR; B75494; B75494.
DR TIGR; DR0653; -.
DR GO; GO:0008080; F.N-acetyltransferase activity; IEA.
DR InterPro; IPR000182; GCN5acetyl_trans.
DR InterPro; IPR001092; HLH basic.
DR Pfam; PF00563; Acetyltransf; 1.
DR PROSITE; PS00038; HLH 1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 176 AA; 19099 MW; 1B4A7EC9E1D5BC3B CRC64;

Query Match 65.6%; Score 40; DB 16; Length 176;
Best Local Similarity 61.5%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
Db 144 RASERVLEKIGFA 156

RESULT 24
Q83132 PRELIMINARY; PRT; 502 AA.
AC Q83132;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Putative aminopeptidase.
GN TW261.
OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=218496;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22495039; PubMed=12606174;
RA Bentley S.D., Mairwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,
RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
RA Bartell B.G., Parkhill J., Relman D.A.;
RT "Sequencing and analysis of the genome of the Whipple's disease
RT bacterium Tropheryma whipplei.";
RL Lancet 361:637-644 (2003).
DR EMBL; BX251410; CAD6937.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004177; P:aminopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008330; Ctsl_aminptdase_B.
DR InterPro; IPR000819; Peptidase_M17_N.
DR InterPro; IPR008283; Peptidase_M17_N.
DR Pfam; PF00883; Peptidase_M17_N; 1.
DR 2fam; PF02789; Peptidase_M17_N; 1.
DR PRINTS; PR00481; LAVNOPPTDASE.
DR PROSITE; PS00631; CYTOSOL_AP; 1.
DR PIRSF; PIRSF036388; Ctsl_aminptdase_B; 1.
KW Aminopeptidase; Complete proteome.
SQ SEQUENCE 502 AA; 53336 MW; 5249BEC6BD06C73 CRC64;

Query Match 63.9%; Score 39; DB 16; Length 502;
Best Local Similarity 72.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 11
Db 42 RAYEDILQKLG 52

RESULT 25
Q97PE5 PRELIMINARY; PRT; 505 AA.
AC Q97PE5;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE PTS system, IIBC components.
GN SPI684.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21337209; PubMed=11463916;
RA Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Mayam L.A., White O., Salzberg S.L., Lewis M.R., Radure D.,
RA Holtzapflee E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506 (2001).
DR EMBL; AE007461; AAK75763.1; -.
DR PIR; B95196; B95196.
DR TIGR; SPI684; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; P:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR003352; Ptrans_EIIC.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR TIGRfams; TIGR00826; EIIB_glc; 1.

```

```

KW Complete proteome.
SQ SEQUENCE 505 AA; 54470 MW; 586A20E3C6DAD89E CRC64;

Query Match 63.9%; Score 39; DB 16; Length 505;
Best Local Similarity 88.9%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLGKA 13
DB 3 KVLQKVGKA 11

RESULT 26
Q8DNV7 PRELIMINARY; PRT; 510 AA.
AC Q8DNV7;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Phosphotransferase system sugar-specific EII component (EC 2.7.1.69).
GN PTS-EII OR SPR1528.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rokey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Kosteck P.R., Jr., Skatrud P.L.,
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RT J. Bacteriol. 183:5709-5717(2001).
RL EMBL; AE008521; AAL00332.1; -.
DR PIR; G98062; G98062.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho...; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001996; P:trans_EIIB.
DR InterPro; IPR003352; P:trans_EIIC.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR TIGRFAMs; TIGR00826; EIIB_glc; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 510 AA; 55033 MW; 5A34BA199F15DF33 CRC64;

Query Match 63.9%; Score 39; DB 16; Length 510;
Best Local Similarity 88.9%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLGKA 13
DB 8 KVLQKVGKA 16

RESULT 27
Q8RBN7 PRELIMINARY; PRT; 513 AA.
AC Q8RBN7;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein TTE0779.
GN TTE0779.

Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=2192816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
BL Genome Res. 12:689-700(2002).
DR EMBL; AE013045; AAM24036.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 513 AA; 56120 MW; 3B91AE6C63402894 CRC64;

Query Match 63.9%; Score 39; DB 16; Length 513;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12
DB 388 RIREKILETKGK 399

RESULT 28
Q42701 PRELIMINARY; PRT; 516 AA.
AC Q42701;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cytochrome P450 (Fragment).
GN CYP72C.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Gentianales; Apocynaceae; Rauvolfioideae; Vinaceae;
OC Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cp3;
RA Mangold U., Eichel J., Batschauer A., Lanz T., Kaiser T.,
RA Spangenberg G., Werck-Reichhart D., Schroeder J.;
RT "Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from
RT Catharanthus roseus, and transgenic expression in tobacco and
RT Arabidopsis thaliana.";
RL Plant Sci. 96:129-136(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Cp3;
RA Joachim Schroeder;
RA Submitted (SEP-1993) to the EMBL/GenBank/DDBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; L19075; AAA17745.1; -.
DR PIR; T10000; T10000.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR011128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON TER 1
SQ SEQUENCE 516 AA; 59720 MW; 02A7D9B0936D931P CRC64;

Query Match 63.9%; Score 39; DB 10; Length 516;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12

```



## RESULT 32

Q9BS66 ID Q9BS66 PRELIMINARY; PRT; 88 AA.  
 AC Q9BS66  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Similar to epilepsy, progressive myoclonus type 2, Lafora disease (Laforin).  
 DE Laforin.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC005286; AAH05286.1; -.  
 DR GO; GO:0004721; P:protein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
 SQ SEQUENCE 88 AA; 9933 MW; D28FAB18CC28SD07 CRC64;

Query Match 62.3%; Score 38; DB 4; Length 88;  
 Best Local Similarity 66.7%; Pred. No. 53;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
 Db 69 RAQEDFFQKFGK 80  
 |||||  
 |||||

## RESULT 33

Q8IU96 ID Q8IU96 PRELIMINARY; PRT; 193 AA.  
 AC Q8IU96  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Laforin.  
 GN EPM2A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Ganesh S., Yamakawa K.;  
 RL "Cloning of differentially spliced transcripts of the EPM2A gene."  
 DR EMBL; AF454492; AA015524.1; -.  
 DR EMBL; AF454493; AA015525.1; -.  
 DR GO; GO:0008138; P:protein tyrosine/serine/threonine phosphatase...; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR000340; DS\_phosphatase.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF00782; DSPC; 1.  
 DR SMART; SM00195; DSPC; 1.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 SQ SEQUENCE 193 AA; 22160 MW; 3DC9436A9885B915 CRC64;

Query Match 62.3%; Score 38; DB 4; Length 193;  
 Best Local Similarity 66.7%; Pred. No. 11e+02;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
 Db 174 RAQEDFFQKFGK 185  
 |||||  
 |||||

## RESULT 34

Q9UEN2 ID Q9UEN2 PRELIMINARY; PRT; 250 AA.  
 AC Q9UEN2  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE LAPTIFase (Fragment).  
 DE LAPTIFase (Fragment).  
 GN EPM2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Serracosa J.M., Gomez-Garre P., Gallardo E., Anta B.,  
 RA Beltran-Valero de Bernabe D., Lindhout D., Augustin P.B.,  
 RA Tassinari C., Michelucci R., Malafosse A., Topcu M., Grid D.,  
 RA Dravet C., Berkovic S., Rodriguez de Cordoba S.;  
 RT "A novel protein tyrosine phosphatase gene is mutated in progressive myoclonus epilepsy of the Lafora type (EPM2).";  
 RL Hum. Mol. Genet. 8:0-0(1999).  
 DR EMBL; AJ130763; CAA10199.1; -.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0008138; P:protein tyrosine/serine/threonine phosphatase...; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR002044; CBD\_4.  
 DR InterPro; IPR000340; DS\_phosphatase.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF00782; DSPC; 1.  
 DR PRODOM; PD001568; CBD\_4; 1.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 FT NON TER  
 SQ SEQUENCE 250 AA; 28750 MW; 6AA1913D31F2A8E3 CRC64;

Query Match 62.3%; Score 38; DB 4; Length 250;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12  
 Db 231 RAQEDFFQKFGK 242  
 |||||  
 |||||

## RESULT 35

Q8IX25 ID Q8IX25 PRELIMINARY; PRT; 257 AA.  
 AC Q8IX25  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Laforin (Fragment).  
 GN EPM2A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Ganesh S., Yamakawa K.;  
 RL "Cloning of differentially spliced transcripts of the EPM2A gene."  
 DR EMBL; AF454491; AA015523.1; -.  
 DR GO; GO:0008138; P:protein tyrosine/serine/threonine phosphatase...; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR000340; DS\_phosphatase.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR SMART; SM00195; DSPC; 1.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.

```

DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
FT NON_TER
SQ SEQUENCE 257 AA; 29473 MW; BB3BEEB5C9E1E4CE CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 4; Length 257;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 12
DB 238 RAQEDFFQKFGK 249

RESULT 36
Q8TT25 PRELIMINARY; PRT; 300 AA.
AC Q8TT25;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 3-hydroxyisobutyrate dehydrogenase.
GN MA0614.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
CX NCBI_TaxID=2214;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nuebaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S.C., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., Mcowan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Fritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010722; BAM04058.1;
DR GO; GO:0004616; P:phosphogluconate dehydrogenase (decarboxyla. . .; IEA.
DR GO; GO:0006098; P:pentose-phosphate shunt; IEA.
DR InterPro; IPR006183; 6PGD.
DR Pfam; PF03446; NAD binding.2; 1.
DR PRINTS; PR00076; 6PGDHRGNASE.
KW Complete proteome.
SQ SEQUENCE 300 AA; 32287 MW; FB1B7A8A98A8E8A7E CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 17; Length 300;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGKA 13
DB 153 RIKHFLNKLKGS 165

RESULT 37
O95278 PRELIMINARY; PRT; 331 AA.
AC O95278;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LAFORIN (PROGRESSIVE MYOCLONUS epilepsy type 2).
GN EPM2A.
OS Homo sapiens (Human).

```

---

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP MEDLINE=98442653; PubMed=9771710;
RX Minassian B.A., Lee J.R., Herbrick J.A., HuiZeng J., Soder S.,
RA Mugali A.J., Dunham I., Gardner R., Fong C.Y., Carpenter S.,
RA Jardim L., Satishchandra P., Andermann E., Shead O.C. III,
RA Lopes-Cendes I., Tsui L.C., Delgado-Escueta A.V., Rouleau G.A.,
RA Scherer S.W.;
RT "Mutations in a gene encoding a novel protein tyrosine phosphatase
RT cause progressive myoclonus epilepsy.";
RL Nat. Genet. 20:171-174(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lee J.R., Scherer S.W.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Minassian B.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11001928;
RA Ganesh S., Agarwala K.I., Ueda K., Akagi T., Shoda K., Usui T.,
RA Hashikawa T., Osada H., Delgado-Escueta A.V., Yamakawa K.;
RT "Laforin, defective in the progressive myoclonus epilepsy of Lafora
RT type, is a dual-specificity phosphatase associated with
RT polyribosomes.";
RL Hum. Mol. Genet. 9:2251-2261(2000).
DR EMBL; AF084535; AAC83347.2; -.
DR EMBL; AF284580; AAG18377.1; -.
DR Genbank; HGNC:3413; EPM2A.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0008138; P:protein tyrosine/serine/threonine phosphatase. . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR002044; CBD_4.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000340; DS_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR ProDom; PD001568; CBD_4; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 331 AA; 37157 MW; DD79F917262AB458 CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 4; Length 331;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 12
DB 312 RAQEDFFQKFGK 323

RESULT 38
Q7T2D2 PRELIMINARY; PRT; 408 AA.
ID Q7T2D2
AC Q7T2D2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22386257; PubMed=12477932;

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young J.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Touchman J.W., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX STRAIN=305 / ATCC 13902 / XV 101;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorfy H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formigueri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL; AE011883; AM37312.1; -;  
 DR InterPro; IPR006463; M1AB\_methiolase.  
 DR InterPro; IPR007197; Radical SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR TIGRFAMs; TIGR01574; m1AB-methiolase; 1.  
 DR TIGRFAMs; TIGR00089; TIGR00089; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 452 AA; 49662 MW; 0DSBEAB175C4PB3E CRC64;

Query Match 62.3%; Score 38; DB 13; Length 408;  
 Best Local Similarity 58.3%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 ID Q89AC2 PRELIMINARY; PRT; 445 AA.  
 AC Q89AC2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN YLEA OR BFP392.  
 OS Buchnera aphidicola (subsp. Baizongia pistaciae).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=135842;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=22425901; PubMed=12522265;  
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,  
 RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,  
 RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;  
 RT "Reductive genome evolution in Buchnera aphidicola.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).  
 DR EMBL; AE014017; AAC27104.1; -;  
 DR InterPro; IPR007197; Radical SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 445 AA; 51265 MW; 5B5F5BF3BCEIDE CRC64;

Query Match 62.3%; Score 38; DB 16; Length 445;  
 Best Local Similarity 58.3%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 DB 100 RLQERLLKLGK 111  
 :|||||:|:  
 :|:|:|:|:|:  
 :|:|:|:|:|:

RESULT 39  
 Q89AC2 PRELIMINARY; PRT; 445 AA.  
 ID Q89AC2;  
 AC Q89AC2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN YLEA OR BFP392.  
 OS Buchnera aphidicola (subsp. Baizongia pistaciae).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=135842;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=22425901; PubMed=12522265;  
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,  
 RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,  
 RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;  
 RT "Reductive genome evolution in Buchnera aphidicola.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).  
 DR EMBL; AE014017; AAC27104.1; -;  
 DR InterPro; IPR007197; Radical SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 445 AA; 51265 MW; 5B5F5BF3BCEIDE CRC64;

Query Match 62.3%; Score 38; DB 16; Length 445;  
 Best Local Similarity 58.3%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 DB 55 KAQEKVFLQGR 66  
 :|||||:|:  
 :|:|:|:|:|:  
 :|:|:|:|:|:

RESULT 40  
 Q89P99 PRELIMINARY; PRT; 452 AA.  
 ID Q89P99;  
 AC Q89P99;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein XAC2461.  
 GN XAC2461.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=305 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorfy H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formigueri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL; AE011883; AM37312.1; -;  
 DR InterPro; IPR006463; M1AB\_methiolase.  
 DR InterPro; IPR007197; Radical SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR TIGRFAMs; TIGR01574; m1AB-methiolase; 1.  
 DR TIGRFAMs; TIGR00089; TIGR00089; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 452 AA; 49662 MW; 0DSBEAB175C4PB3E CRC64;

Query Match 62.3%; Score 38; DB 16; Length 452;  
 Best Local Similarity 58.3%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 DB 55 KAQEKVFLQGR 66  
 :|||||:|:  
 :|:|:|:|:|:  
 :|:|:~|:|:|:

RESULT 41  
 Q21004 PRELIMINARY; PRT; 461 AA.  
 ID Q21004;  
 AC Q21004;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE F58G6.1 protein.  
 GN F58G6.1.



OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabdityidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lloyd C.R.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 CC 1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; Z68217; CAA92465.1; -;  
 DR PIR; T22946; T22946.  
 DR WormPep; F58G6.1; CE03428.  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PRO1251; AMPHIPHYSIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR SH3 domain.  
 KW SH3 domain.  
 SQ SEQUENCE 461 AA; 51910 MW; 771C99E21EF7BEC8 CRC64;  
 Query Match 62.3%; Score 38; DB 5; Length 461;  
 Best Local Similarity 53.8%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGK 13  
 DB 15 RTKEKLEGIGKA 27  
 RESULT 42  
 QSKTEO PRELIMINARY; PRT; 474 AA.  
 ID Q9KTE0  
 AC Q9KTE0  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein VCO962.  
 GN VCO962.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.P., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Emmlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004177; AAF94124.1; -;  
 DR PIR; E82261; E82261.

TIGR; VC0962;  
 DR InterPro; IPR006638; Elp3.  
 DR InterPro; IPR006463; MiaB\_methylolase.  
 DR InterPro; IPR007197; Radical\_SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical\_SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR SMART; SM00729; Elp3; 1.  
 DR TIGRFAMs; TIGR01574; miaB-methylolase; 1.  
 DR TIGRFAMs; TIGR00089; TIGR00089; 1.  
 DR PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 474 AA; 53613 MW; C1C5DF73CC96C467 CRC64;  
 Query Match 62.3%; Score 38; DB 16; Length 474;  
 Best Local Similarity 59.3%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGK 12  
 DB 54 KAQEKVFHQLGR 65  
 RESULT 43  
 Q9RCI2 PRELIMINARY; PRT; 474 AA.  
 ID Q9RCI2  
 AC Q9RCI2  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE MiaB protein (Methylthiolation of isopentenylated A37 derivatives in  
 DE rRNA).  
 GN MIAB OR STM0670.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20042346; PubMed=10572129;  
 RA Esberg B., Leung H.C.E., Tsui H.C.T., Bjoerk G.R., Winkler M.E.;  
 RT "Identification of the miaB gene involved in the methylthiolation of  
 RT isopentenylated A37 derivatives in the trna of Salmonella typhimurium  
 RT and Escherichia coli.";  
 RL J. Bacteriol. 181:7256-7265(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RT LT2.";  
 RL Nature 413:852-856(2001).  
 DR EMBL; AJ249116; CAB62263.1; -;  
 DR EMBL; AE008727; AAL19621.1; -;  
 DR InterPro; IPR006638; Elp3.  
 DR InterPro; IPR006463; MiaB\_methylolase.  
 DR InterPro; IPR007197; Radical\_SAM.  
 DR InterPro; IPR002792; TRAM.  
 DR InterPro; IPR005839; UPF0004.  
 DR Pfam; PF04055; Radical\_SAM; 1.  
 DR Pfam; PF01938; TRAM; 1.  
 DR Pfam; PF00919; UPF0004; 1.  
 DR SMART; SM00729; Elp3; 1.  
 DR TIGRFAMs; TIGR01574; miaB-methylolase; 1.  
 DR TIGRFAMs; TIGR00089; TIGR00089; 1.  
 DR PROSITE; PS01278; UPF0004; 1.

KW Complete proteome.  
SQ SEQUENCE 474 AA; 53716 MW; 25E5CED5F28D38C CRC64;

Query Match 62.3%; Score 38; DB 16; Length 474;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
:|||||:|:  
DB 54 KAQEKVFHQLGR 65

RESULT 44  
QSDFE8 PRELIMINARY; PRT; 474 AA.  
AC QSDFE8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE 2-methylthioadenine synthetase.  
GN VV10266.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMCP6;  
RA Rhee J.H.; Kim S.Y.; Chung S.S.; Kim J.J.; Moon Y.H.; Jeong H.;  
RA Choy H.E.;  
RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016797; AAC08800.1;  
DR InterPro; IPR006638; Elp3.  
DR InterPro; IPR006463; MiaB\_methylolase.  
DR InterPro; IPR007197; Radical\_SAM.  
DR InterPro; IPR002792; TRAM.  
DR InterPro; IPR005839; UPF0004.  
DR Pfam; PF04055; Radical\_SAM; 1.  
DR Pfam; PF01938; TRAM; 1.  
DR Pfam; PF00919; UPF0004; 1.  
DR SMART; SM00729; Elp3; 1.  
DR TIGSFAMS; TIGR01574; miaB\_methylolase; 1.  
DR TIGSFAMS; TIGR00089; TIGR00089; 1.  
DR PROSITE; PS01278; UPF0004; 1.  
DR Complete proteome.  
KW Complete proteome.  
SQ SEQUENCE 474 AA; 53592 MW; 4B2BC7B749B689A2 CRC64;

Query Match 62.3%; Score 38; DB 16; Length 474;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
:|||||:|:  
DB 54 KAQEKVFHQLGR 65

RESULT 45  
QSDX45 PRELIMINARY; PRT; 474 AA.  
AC QSDX45;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE TRNA-i(6)A37 modification enzyme MiaB.  
GN MIAB OR S01181.  
OS Shewanella oneidensis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
OC Alteromonadaceae; Shewanella.  
OX NCBI\_TaxID=70863;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MR-1;

RX MEDLINE=22297686; PubMed=12368813;  
RA Heidelberg J.P.; Paulsen I.T.; Nelson K.E.; Gaidos E.J.; Nelson W.C.;  
RA Meyer T.D.; Eisen J.A.; Seshadri R.; Ward N.; Methe B.; Clayton R.A.;  
RA Meyer R.T.; Tsapin A.; Scott J.; Beanan M.; Brinkac L.; Daugherty S.;  
RA DeBoy R.T.; Dodson R.J.; Durkin A.S.; Haft D.H.; Kolonay J.F.;  
RA Madupu R.; Peterson J.B.; Umayam L.A.; White O.; Wolf A.M.;  
RA Vamathevan J.; Weidman J.; Imbraim M.; Lee K.; Berry K.; Lee C.;  
RA Mueller J.; Khouri H.; Gill J.; Utterback T.R.; McDonald L.A.;  
RA Feldblyum T.V.; Smith H.O.; Venter J.C.; Neallson K.H.; Fraser C.M.;  
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
*Shewanella oneidensis*.";  
RL Nat. Biotechnol. 20:1118-1123 (2002).  
DR EMBL; AE015582; AAN54251.1;  
DR TIGR; S01181;  
DR InterPro; IPR006638; Elp3.  
DR InterPro; IPR006463; MiaB\_methylolase.  
DR InterPro; IPR007197; Radical\_SAM.  
DR InterPro; IPR002792; TRAM.  
DR InterPro; IPR005839; UPF0004.  
DR Pfam; PF04055; Radical\_SAM; 1.  
DR Pfam; PF01938; TRAM; 1.  
DR Pfam; PF00919; UPF0004; 1.  
DR SMART; SM00729; Elp3; 1.  
DR TIGSFAMS; TIGR01574; miaB\_methylolase; 1.  
DR TIGSFAMS; TIGR00089; TIGR00089; 1.  
DR PROSITE; PS01278; UPF0004; 1.  
DR Complete proteome.  
KW Complete proteome.  
SQ SEQUENCE 474 AA; 53760 MW; 976DCE04F6091BD2 CRC64;

Query Match 62.3%; Score 38; DB 16; Length 474;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
:|||||:|:  
DB 54 KAQEKVFHQLGR 65

RESULT 46  
QSD7R24 PRELIMINARY; PRT; 474 AA.  
ID QSD7R24;  
AC QSD7R24;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Conserved hypothetical protein.  
GN VP0733.  
OS *Vibrio parahaemolyticus*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RMD 2210633 / Serotype O3:K6;  
RA MEDLINE=22508454; PubMed=12820739;  
RA Makino K.; Oshima K.; Kurokawa K.; Yokoyama K.; Uda T.; Tagomori K.;  
RA Iijima Y.; Najima M.; Nakano M.; Yamashita A.; Kubota Y.; Kimura S.;  
RA Yasunaga T.; Honda T.; Shinagawa H.; Hattori M.; Iida T.;  
RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism  
distinct from that of *V. cholerae*.";  
RL Lancet 361:743-749 (2003).  
DR EMBL; AP005075; BAC58996.1;  
DR InterPro; IPR007197; Radical\_SAM.  
DR InterPro; IPR002792; TRAM.  
DR InterPro; IPR005839; UPF0004.  
DR Pfam; PF04055; Radical\_SAM; 1.  
DR Pfam; PF01938; TRAM; 1.  
DR Pfam; PF00919; UPF0004; 1.  
DR PROSITE; PS01278; UPF0004; 1.  
DR Hypothetical protein; Complete proteome.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 474 AA; 53713 MW; 0CA152B9FB75A657 CRC64;

Query Match 62.3%; Score 38; DB 16; Length 474;

```

Best Local Similarity 58.3%; Pred. No. 2.7e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12
DB 54 KAQEKVFHOLGR 65

RESULT 47
Q83LY3
ID Q83LY3 PRELIMINARY; PRT; 474 AA.
AC
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orf, conserved hypothetical protein (Hypothetical protein ylea).
GN YLEA OR SF0621 OR S0643.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE015093; AAN42258.1; -
DR EMBL; AE015980; AAP16128.1; -
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR006463; MlaB_methylolase.
DR InterPro; IPR007197; Radical_SAM.
DR InterPro; IPR002792; TRAM.
DR InterPro; IPR005839; UPF0004.
DR Pfam; PF04055; Radical_SAM; 1.
DR Pfam; PF01938; TRAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR01574; mlaB-methylolase; 1.
DR TIGRFAMs; TIGR00089; TIGR00089; 1.
DR PROSITE; PS01278; UPF0004; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 474 AA; 53695 MW; 537A6093888F7F9D CRC64;

Query Match 62.3%; Score 38; DB 16; Length 474;
Best Local Similarity 58.3%; Pred. No. 2.7e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12
DB 54 KAQEKVFHOLGR 65

RESULT 48
Q87AP4
ID Q87AP4 PRELIMINARY; PRT; 475 AA.

```

```

Q87AP4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN PD1779.
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Cluyt M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Silva F.R.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Perro M.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuranae E.E.,
RA Martino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.P., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.B. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RL J. Bacteriol. 185:1018-1026 (2003).
DR EMBL; AE012559; AAO29613.1; -
DR InterPro; IPR007197; Radical_SAM.
DR InterPro; IPR002792; TRAM.
DR InterPro; IPR005839; UPF0004.
DR Pfam; PF04055; Radical_SAM; 1.
DR Pfam; PF01938; TRAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR PROSITE; PS01278; UPF0004; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 475 AA; 53035 MW; 84C60B43B197B9DC CRC64;

Query Match 62.3%; Score 38; DB 16; Length 475;
Best Local Similarity 58.3%; Pred. No. 2.7e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12
DB 77 KAQEKVFHOLGR 88

RESULT 49
Q7VP74
ID Q7VP74 PRELIMINARY; PRT; 475 AA.
AC Q7VP74;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN HD0226.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.,
RA "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017151; AAP95213.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 475 AA; 53638 MW; CB0F3B354C5F2DFA CRC64;

```

Query Match 62.3%; Score 38; DB 16; Length 475;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
      :|||||:|  
DB 53 KAQEKVFSQGLR 64

## RESULT 50

Q8P8B5  
ID Q8P8B5 PRELIMINARY; PRT; 485 AA.  
AC Q8P8B5;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Hypothetical protein XCC2328.  
GN XCC2328.  
OS Xanthomonas campestris (pv. campestris).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=340;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33913 / NCPPB 528;  
RX MEDLINE=22022145; PubMed=12024217;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,  
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
RA Paria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos E.G.M., Lemos M.V.F.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Locali E.C., Machado M.A., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spinoia L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities."  
RL Nature 417:459-463(2002).  
DR EMEL; AB012340; AAM41606.1;  
DR InterPro; IPR006463; MiaB\_methylase.  
DR InterPro; IPR007197; Radical\_SAM.  
DR InterPro; IPR002792; TRAM.  
DR InterPro; IPR005839; UPF0004.  
DR Pfam; PF04055; Radical\_SAM; 1.  
DR Pfam; PF01938; TRAM; 1.  
DR Pfam; PF00919; UPF0004; 1.  
DR TIGRFAMs; TIGR01574; miaB\_methylase; 1.  
DR TIGRFAMs; TIGR00089; TIGR00089; 1.  
DR PROSITE; PS01278; UPF0004; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 485 AA; 52737 MW; 08C9EA7E10528FEE CRC64;

Query Match 62.3%; Score 38; DB 16; Length 485;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
      :|||||:|  
DB 88 KAQEKVFSQGLR 99

Search completed: March 4, 2004, 17:46:18  
Job time : 37.6129 secs





MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/10/123,807  
FILING DATE: 16-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/445,247  
FILING DATE: 03-Dec-1999  
APPLICATION NUMBER: US 08/870,126  
FILING DATE: 06-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60DPCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-123-807-4

Query Match 100.0%; Score 61; DB 14; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
DB 21 RAQEKVLQKLGKA 33  
|||||

## RESULT 3

US-10-116-275-238  
; Sequence 238, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 238  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-238

Query Match 100.0%; Score 61; DB 15; Length 482;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
DB 24 RAQEKVLQKLGKA 36  
|||||

## RESULT 4

US-10-091-007-198

; Sequence 198, Application US/10091007  
; Publication No. US20030170782A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Le Page, Richard W F  
; APPLICANT: Hanniffy, Sean B  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21978WO  
; CURRENT APPLICATION NUMBER: US/10/091,007  
; CURRENT FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: GB 9921125.2  
; PRIOR FILING DATE: 1999-09-07  
; NUMBER OF SEQ ID NOS: 276  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 198  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Streptococcus agalactiae  
US-10-091-007-198

Query Match 67.2%; Score 41; DB 14; Length 373;  
Best Local Similarity 63.6%; Pred. No. 0.47;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13  
DB 337 EEKILKLGKS 347  
|||||

## RESULT 5

US-09-769-787-120  
; Sequence 120, Application US/09769787  
; Publication No. US20030091577A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Gilbert, Christophe FG  
; APPLICANT: Hansbro, Philip M  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21129WO  
; CURRENT APPLICATION NUMBER: US/09/769,787  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: GB 9816337.1  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: US 60/125164  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 388  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 120  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-769-787-120

Query Match 63.9%; Score 39; DB 10; Length 505;  
Best Local Similarity 88.9%; Pred. No. 1.4e+02;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLGKA 13  
DB 3 KVLQKVGKA 11  
|||||

## RESULT 6

US-03-815-242-13474  
; Sequence 13474, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zvekind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13474  
LENGTH: 510  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13474

Query Match 63.9%; Score 39; DB 9; Length 510;  
Best Local Similarity 88.9%; Pred. No. 1.4e-02;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLGKA 13  
|||:|  
DB 8 KVLQKVGKA 16

RESULT 7  
US-09-815-242-13607  
Sequence 13607, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13607

LENGTH: 510  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13607  
Query Match 63.9%; Score 39; DB 9; Length 510;  
Best Local Similarity 88.9%; Pred. No. 1.4e-02;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KVLQKLGKA 13  
|||:|  
DB 8 KVLQKVGKA 16  
RESULT 8  
US-10-369-493-13329  
Sequence 13329, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 13329  
LENGTH: 320  
TYPE: PRT  
ORGANISM: Aspergillus nidulans  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(320)  
OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-13329

Query Match 62.3%; Score 38; DB 15; Length 320;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
|||:|  
DB 2 RQYRVLQQLGK 13

RESULT 9  
US-10-161-051-57  
Sequence 57, Application US/10161051  
Publication No. US20030152945A1  
GENERAL INFORMATION:  
APPLICANT: Peter Deak  
APPLICANT: David M Glover  
APPLICANT: Carol Midgley  
TITLE OF INVENTION: Cell cycle progression proteins  
FILE REFERENCE: CCI-021CP  
CURRENT APPLICATION NUMBER: US/10/161,051  
CURRENT FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: GB 0007268.6  
PRIOR FILING DATE: 2000-03-24  
NUMBER OF SEQ ID NOS: 194  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 57  
LENGTH: 357  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-10-161-051-57



Query Match 62.3%; Score 38; DB 14; Length 357;  
Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLGKA 13  
|:|:|:|:|:  
Db 29 QKQMLEKLGKA 39

## RESULT 10

US-10-002-278-2  
; Sequence 2, Application US/10002278  
; Publication No. US20020132234A1  
; GENERAL INFORMATION:  
; APPLICANT: Jessell, Thomas M.  
; APPLICANT: Basler, Konard  
; APPLICANT: Yamada, Toshiya  
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF DORSALIN-1  
; FILE REFERENCE: 0575/40314-A  
; CURRENT APPLICATION NUMBER: US/10/002,278  
; CURRENT FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 427  
; TYPE: PRT  
; ORGANISM: Chick  
US-10-002-278-2

Query Match 62.3%; Score 38; DB 13; Length 427;  
Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGK 12  
|:|:|:|:|:  
Db 282 QBSVLNKLK 291

## RESULT 11

US-09-864-761-47746  
; Sequence 47746, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Ranzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmiga-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47746  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC009640.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
; OTHER INFORMATION: EST\_HUMAN HIT: BF692371.1, EVALUE 2.00e-08  
US-09-864-761-47746

Query Match 60.7%; Score 37; DB 9; Length 33;  
Best Local Similarity 53.8%; Pred. No. 19;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
|:|:|:|:|:  
Db 3 QAERKILKLGKA 15

## RESULT 12

US-09-738-973-182  
; Sequence 182, Application US/09738973  
; Patent No. US20020110563A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Fling, Steven P.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Indirias, Carol Yoseph  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Elliot, Mark  
; APPLICANT: Mannion, Jane  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 21021.475C9  
; CURRENT APPLICATION NUMBER: US/09/738,973  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 587  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 182  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-738-973-182

Query Match 60.7%; Score 37; DB 9; Length 209;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKIG 11  
|:|:|:|:|:  
Db 141 AREKVLQTLG 150

```
RESULT 13
US-09-854-133-182
; Sequence 182, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 182
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-182

Query Match      60.7%; Score 37; DB 9; Length 209;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11
|:|||||
Db 141 AREKVLQTLG 150

RESULT 14
US-10-144-649A-182
; Sequence 182, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 182
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-144-649A-182

Query Match      60.7%; Score 37; DB 14; Length 209;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11
|:|||||
Db 141 AREKVLQTLG 150

RESULT 15
US-10-369-493-1040
; Sequence 1040, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1040
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-10-369-493-1040

Query Match      60.7%; Score 37; DB 15; Length 375;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EKVQLKLGKA 13
|:|||||
Db 248 EKELKLGKA 257

RESULT 16
US-10-369-493-1103
; Sequence 1103, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1103
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
US-10-369-493-1103

Query Match      60.7%; Score 37; DB 15; Length 403;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGK 12
|:|||||
Db 362 QGKTLQKLGK 371

RESULT 17
US-09-933-767-408
; Sequence 408, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
```

; PRIOR APPLICATION NUMBER: 60/193,170  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: 09/205,258  
 ; PRIOR FILING DATE: 1998-12-04  
 ; PRIOR APPLICATION NUMBER: PCT/US98/11422  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/048,885  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/049,375  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,881  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,880  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,896  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/049,020  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,876  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,895  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,884  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,894  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,971  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,964  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,882  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,899  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,893  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,900  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,901  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,922  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,915  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/049,019  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,970  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,972  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,916  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/049,373  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,875  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/049,374  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,917  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,949  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,374  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,883  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,897  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,898  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,962  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,963

; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,877  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/048,878  
 ; PRIOR FILING DATE: 1997-06-06  
 ; PRIOR APPLICATION NUMBER: 60/068,054  
 ; PRIOR FILING DATE: 1997-12-18  
 ; PRIOR APPLICATION NUMBER: 60/068,064  
 ; PRIOR FILING DATE: 1997-12-18  
 ; PRIOR APPLICATION NUMBER: 60/068,053  
 ; PRIOR FILING DATE: 1997-12-18  
 ; PRIOR APPLICATION NUMBER: 60/070,923  
 ; PRIOR FILING DATE: 1997-12-18  
 ; PRIOR APPLICATION NUMBER: 60/073,160  
 ; PRIOR FILING DATE: 1998-01-30  
 ; PRIOR APPLICATION NUMBER: 60/073,159  
 ; PRIOR FILING DATE: 1998-01-30  
 ; PRIOR APPLICATION NUMBER: 60/073,165  
 ; PRIOR FILING DATE: 1998-01-30  
 ; PRIOR APPLICATION NUMBER: 60/073,164  
 ; PRIOR FILING DATE: 1998-01-30  
 ; PRIOR APPLICATION NUMBER: 60/085,925  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,921  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,923  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,922  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/092,921  
 ; PRIOR FILING DATE: 1998-07-15  
 ; PRIOR APPLICATION NUMBER: 60/094,657  
 ; PRIOR FILING DATE: 1998-07-30  
 ; NUMBER OF SEQ ID NOS: 1245  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 408  
 ; LENGTH: 447  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (447)  
 ; OTHER INFORMATION: Xaa equals stop translation  
 US-09-933-767-408  
 Query Match 60.7%; Score 37; DB 10; Length 447;  
 Best Local Similarity 80.0%; Pred. No. 2.7e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 AQEKVLQKLG 11  
 Db 378 AREKVLQTLG 387  
 RESULT 18  
 US-10-023-282-408  
 ; Sequence 408, Application US/10023282  
 ; Publication No. US20030092893A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young et al.  
 ; TITLE OF INVENTION: 207 Human Secreted Proteins  
 ; FILE REFERENCE: PZ007P1  
 ; CURRENT APPLICATION NUMBER: US/10/023,282  
 ; CURRENT FILING DATE: 2001-12-20  
 ; EARLIER APPLICATION NUMBER: 09/205,258  
 ; EARLIER FILING DATE: 1998-12-04  
 ; EARLIER APPLICATION NUMBER: PCT/US98/11422  
 ; EARLIER FILING DATE: 1998-06-04  
 ; EARLIER APPLICATION NUMBER: 60/048,885  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,375  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,881

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30

```

```

; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 408
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (447)
; OTHER INFORMATION: Xaa equals stop translation
US-10-023-282-408

```

```

Query Match      60.7%; Score 37; DB 14; Length 447;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 AOEKVLQKLG 11
   :|:|:|:|:|
Db 378 AREKVLQTLG 387

```

RESULT 19

```

US-09-935-390A-26
; Sequence 26, Application US/09935390A
; Patent No. US20020076761A1
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; Quianjin, Hu
; Garcia, Pablo
; Williams, Lewis T.
; Kothakota, Srinivas

```

```

TITLE OF INVENTION: Secreted Human Proteins
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/935,390A
FILING DATE: 22-Aug-2001
CLASSIFICATION: <Unknown>

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/988,671
FILING DATE: 1997-12-11
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. R. Potter
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 1369.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2718
TELEFAX: (510) 655-3542
TELEX: <Unknown>

```

```

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. US20020076761A1
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-935-390A-26

```

```

Query Match      60.7%; Score 37; DB 9; Length 451;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 AOEKVLQKLG 11  
|:|||||  
Db 383 ABEKVLQTLG 392

RESULT 20  
US-09-989-722-301  
; Sequence 301, Application US/09989722  
; Patent No. US20020072067A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashtkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730P1C63  
; CURRENT APPLICATION NUMBER: US/09/989,722  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-15  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22

```

; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

```

Query Match      60.7%; Score 37; DB 9; Length 461;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches      8; Conservative      1; Mismatches      1; Indels      0; Gaps      0;

```

```

QY      2 AOEKVLQKLG 11
      |:|||||
Db      393 AREKVLQTLG 402

```

RESULT 21

```

US-09-989-723-301
; Sequence 301. Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1062
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04

```

; PRIOR APPLICATION NUMBER: 60/088033  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088326  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088167  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088202  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088212  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088217  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088655  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: 60/088734  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088738  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088742  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088810  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088858  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-15  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355

; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 0; Gaps 0;

QY 2 AQEVLQKLG 11  
 Db 393 AREKVLQTLG 402

RESULT 22

US-09-989-279-301  
 ; Sequence 301, Application US/09989279  
 ; Patent No. US20020072496A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerriksen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC56  
CURRENT APPLICATION NUMBER: US/09/989,279  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24



; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match Best Local Similarity 60.7%; Score 37; DB 9; Length 461;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AQEKVLQKLG 11

Db 393 AREKVLQTLG 402

RESULT 23

US-09-989-727-301

Sequence 301, Application US/09989727

Patent No. US20020072497A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; TITLE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2730F1C65  
 ; CURRENT APPLICATION NUMBER: US/09/989,727  
 ; CURRENT FILING DATE: 2001-11-19  
 ; PRIOR APPLICATION NUMBER: 60/049787  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/065186  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066770  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/075945  
 ; PRIOR FILING DATE: 1998-02-25  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/087106  
 ; PRIOR FILING DATE: 1998-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087609  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087759  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087827  
 ; PRIOR FILING DATE: 1998-06-03  
 ; PRIOR APPLICATION NUMBER: 60/088021  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088025  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088026  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088028  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088029  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088030  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088033  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088326  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088167  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088202  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088212  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088217  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088655  
 ; PRIOR FILING DATE: 1998-06-09

; PRIOR APPLICATION NUMBER: 60/088734  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088738  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088742  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088810  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088858  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088910  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535

; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AOEKVLQKLG 11  
 | : | | | | | |  
 Db 393 AREKVLQTLG 402

RESULT 24  
 US-09-989-731-301  
 ; Sequence 301, Application US/099899731  
 ; Patent No. US20020103125A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.



```

; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY      2 AQRKVLQKLG 11
      :|:|:|:|:|
Db      393 AREKVLQTLG 402

```

```

RESULT 25
US-09-989-732-301
; Sequence 301, Application US/09989732
; Patent No. US20020123463A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```

```

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PLC57
; CURRENT APPLICATION NUMBER: US/09/989,732
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11

```

;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694

;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 AQEKVLQKLG 11  
|:|||||  
DB 393 AREKVLQTLG 402

## RESULT 26

US-09-991-073-301  
; Sequence 301, Application US/09991073  
; Patent No. US20020127576A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurnsey, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC15  
; CURRENT APPLICATION NUMBER: US/09/991, 073  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16

1 PRIOR APPLICATION NUMBER: 60/062250  
2 PRIOR FILING DATE: 1997-10-17  
3 PRIOR APPLICATION NUMBER: 60/065186  
4 PRIOR FILING DATE: 1997-11-12  
5 PRIOR APPLICATION NUMBER: 60/065311  
6 PRIOR FILING DATE: 1997-11-13  
7 PRIOR APPLICATION NUMBER: 60/066770  
8 PRIOR FILING DATE: 1997-11-24  
9 PRIOR APPLICATION NUMBER: 60/075945  
10 PRIOR FILING DATE: 1998-02-25  
11 PRIOR APPLICATION NUMBER: 60/078910  
12 PRIOR FILING DATE: 1998-03-20  
13 PRIOR APPLICATION NUMBER: 60/083322  
14 PRIOR FILING DATE: 1998-04-28  
15 PRIOR APPLICATION NUMBER: 60/084600  
16 PRIOR FILING DATE: 1998-05-07  
17 PRIOR APPLICATION NUMBER: 60/087106  
18 PRIOR FILING DATE: 1998-05-28  
19 PRIOR APPLICATION NUMBER: 60/087607  
20 PRIOR FILING DATE: 1998-06-02  
21 PRIOR APPLICATION NUMBER: 60/087609  
22 PRIOR FILING DATE: 1998-06-02  
23 PRIOR APPLICATION NUMBER: 60/087759  
24 PRIOR FILING DATE: 1998-06-02  
25 PRIOR APPLICATION NUMBER: 60/087827  
26 PRIOR FILING DATE: 1998-06-03  
27 PRIOR APPLICATION NUMBER: 60/088021  
28 PRIOR FILING DATE: 1998-06-04  
29 PRIOR APPLICATION NUMBER: 60/088025  
30 PRIOR FILING DATE: 1998-06-04  
31 PRIOR APPLICATION NUMBER: 60/088026  
32 PRIOR FILING DATE: 1998-06-04  
33 PRIOR APPLICATION NUMBER: 60/088028  
34 PRIOR FILING DATE: 1998-06-04  
35 PRIOR APPLICATION NUMBER: 60/088029  
36 PRIOR FILING DATE: 1998-06-04  
37 PRIOR APPLICATION NUMBER: 60/088030  
38 PRIOR FILING DATE: 1998-06-04  
39 PRIOR APPLICATION NUMBER: 60/088033  
40 PRIOR FILING DATE: 1998-06-04  
41 PRIOR APPLICATION NUMBER: 60/088326  
42 PRIOR FILING DATE: 1998-06-04  
43 PRIOR APPLICATION NUMBER: 60/088167  
44 PRIOR FILING DATE: 1998-06-05  
45 PRIOR APPLICATION NUMBER: 60/088202  
46 PRIOR FILING DATE: 1998-06-05  
47 PRIOR APPLICATION NUMBER: 60/088212  
48 PRIOR FILING DATE: 1998-06-05  
49 PRIOR APPLICATION NUMBER: 60/088217  
50 PRIOR FILING DATE: 1998-06-05  
51 PRIOR APPLICATION NUMBER: 60/088655  
52 PRIOR FILING DATE: 1998-06-09  
53 PRIOR APPLICATION NUMBER: 60/088734  
54 PRIOR FILING DATE: 1998-06-10  
55 PRIOR APPLICATION NUMBER: 60/088738  
56 PRIOR FILING DATE: 1998-06-10  
57 PRIOR APPLICATION NUMBER: 60/088742  
58 PRIOR FILING DATE: 1998-06-10  
59 PRIOR APPLICATION NUMBER: 60/088810  
60 PRIOR FILING DATE: 1998-06-10  
61 PRIOR APPLICATION NUMBER: 60/088824  
62 PRIOR FILING DATE: 1998-06-10  
63 PRIOR APPLICATION NUMBER: 60/088826  
64 PRIOR FILING DATE: 1998-06-10  
65 PRIOR APPLICATION NUMBER: 60/088858  
66 PRIOR FILING DATE: 1998-06-11  
67 PRIOR APPLICATION NUMBER: 60/088861  
68 PRIOR FILING DATE: 1998-06-11  
69 PRIOR APPLICATION NUMBER: 60/088876  
70 PRIOR FILING DATE: 1998-06-11  
71 PRIOR APPLICATION NUMBER: 60/089105  
72 PRIOR FILING DATE: 1998-06-12  
73 PRIOR APPLICATION NUMBER: 60/089440  
74 PRIOR FILING DATE: 1998-06-16  
75 PRIOR APPLICATION NUMBER: 60/089512  
76 PRIOR FILING DATE: 1998-06-16  
77 PRIOR APPLICATION NUMBER: 60/089514  
78 PRIOR FILING DATE: 1998-06-16  
79 PRIOR APPLICATION NUMBER: 60/089532  
80 PRIOR FILING DATE: 1998-06-17  
81 PRIOR APPLICATION NUMBER: 60/089538  
82 PRIOR FILING DATE: 1998-06-17  
83 PRIOR APPLICATION NUMBER: 60/089598  
84 PRIOR FILING DATE: 1998-06-17  
85 PRIOR APPLICATION NUMBER: 60/089599  
86 PRIOR FILING DATE: 1998-06-17  
87 PRIOR APPLICATION NUMBER: 60/089600  
88 PRIOR FILING DATE: 1998-06-17  
89 PRIOR APPLICATION NUMBER: 60/089653  
90 PRIOR FILING DATE: 1998-06-17  
91 PRIOR APPLICATION NUMBER: 60/089801  
92 PRIOR FILING DATE: 1998-06-18  
93 PRIOR APPLICATION NUMBER: 60/089907  
94 PRIOR FILING DATE: 1998-06-18  
95 PRIOR APPLICATION NUMBER: 60/089908  
96 PRIOR FILING DATE: 1998-06-18  
97 PRIOR APPLICATION NUMBER: 60/089947  
98 PRIOR FILING DATE: 1998-06-19  
99 PRIOR APPLICATION NUMBER: 60/089948  
100 PRIOR FILING DATE: 1998-06-19  
101 PRIOR APPLICATION NUMBER: 60/089952  
102 PRIOR FILING DATE: 1998-06-19  
103 PRIOR APPLICATION NUMBER: 60/090246  
104 PRIOR FILING DATE: 1998-06-22  
105 PRIOR APPLICATION NUMBER: 60/090252  
106 PRIOR FILING DATE: 1998-06-22  
107 PRIOR APPLICATION NUMBER: 60/090254  
108 PRIOR FILING DATE: 1998-06-22  
109 PRIOR APPLICATION NUMBER: 60/090349  
110 PRIOR FILING DATE: 1998-06-23  
111 PRIOR APPLICATION NUMBER: 60/090355  
112 PRIOR FILING DATE: 1998-06-23  
113 PRIOR APPLICATION NUMBER: 60/090429  
114 PRIOR FILING DATE: 1998-06-24  
115 PRIOR APPLICATION NUMBER: 60/090431  
116 PRIOR FILING DATE: 1998-06-24  
117 PRIOR APPLICATION NUMBER: 60/090435  
118 PRIOR FILING DATE: 1998-06-24  
119 PRIOR APPLICATION NUMBER: 60/090444  
120 PRIOR FILING DATE: 1998-06-24  
121 PRIOR APPLICATION NUMBER: 60/090445  
122 PRIOR FILING DATE: 1998-06-24  
123 PRIOR APPLICATION NUMBER: 60/090472  
124 PRIOR FILING DATE: 1998-06-24  
125 PRIOR APPLICATION NUMBER: 60/090535  
126 PRIOR FILING DATE: 1998-06-24  
127 PRIOR APPLICATION NUMBER: 60/090540  
128 PRIOR FILING DATE: 1998-06-24  
129 PRIOR APPLICATION NUMBER: 60/090542  
130 PRIOR FILING DATE: 1998-06-24  
131 PRIOR APPLICATION NUMBER: 60/090557  
132 PRIOR FILING DATE: 1998-06-24  
133 PRIOR APPLICATION NUMBER: 60/090676  
134 PRIOR FILING DATE: 1998-06-25  
135 PRIOR APPLICATION NUMBER: 60/090678  
136 PRIOR FILING DATE: 1998-06-25  
137 PRIOR APPLICATION NUMBER: 60/090690  
138 PRIOR FILING DATE: 1998-06-25  
139 PRIOR APPLICATION NUMBER: 60/090694  
140 PRIOR FILING DATE: 1998-06-25  
141 PRIOR APPLICATION NUMBER: 60/090695  
142 PRIOR FILING DATE: 1998-06-25  
143 PRIOR APPLICATION NUMBER: 60/090696  
144 PRIOR FILING DATE: 1998-06-25  
145 PRIOR APPLICATION NUMBER: 60/090862  
146 PRIOR FILING DATE: 1998-06-26

; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ADEKVLQKLG 11  
 DB 393 ADEKVLQTLG 402

## RESULT 27

US-09-990-442-301  
 ; Sequence 301, Application US/09990442  
 ; Patent No. US20020132252A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gertitsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas P.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2730P1C8  
 ; CURRENT APPLICATION NUMBER: US/09/990,442  
 ; CURRENT FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/049787  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/065186  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/075945  
 ; PRIOR FILING DATE: 1998-02-25  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/087106  
 ; PRIOR FILING DATE: 1998-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087609  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087759  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087827  
 ; PRIOR FILING DATE: 1998-06-03  
 ; PRIOR APPLICATION NUMBER: 60/088021  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088025  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088026  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088028  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088029  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088030  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088033  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088326  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088167  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088202  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088212  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088217  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088655  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: 60/088734  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088738  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088742  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088810  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088858  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544

; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11  
 :|||||  
 Db 393 AREKVLQTLG 402

RESULT 28

US-09-991-163-301  
 ; Sequence 301, Application US/09991163  
 ; Patent No. US20020132253A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Klijavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; TITLE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2730P1C17  
 ; CURRENT APPLICATION NUMBER: US/09/991,163  
 ; CURRENT FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/049787  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/065186  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066770  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/075945  
 ; PRIOR FILING DATE: 1998-02-25  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28



1	PRIOR APPLICATION NUMBER: 60/084600
2	PRIOR FILING DATE: 1998-05-07
3	PRIOR APPLICATION NUMBER: 60/087106
4	PRIOR FILING DATE: 1998-05-28
5	PRIOR APPLICATION NUMBER: 60/087607
6	PRIOR FILING DATE: 1998-06-02
7	PRIOR APPLICATION NUMBER: 60/087609
8	PRIOR FILING DATE: 1998-06-02
9	PRIOR APPLICATION NUMBER: 60/087759
10	PRIOR FILING DATE: 1998-06-02
11	PRIOR APPLICATION NUMBER: 60/087827
12	PRIOR FILING DATE: 1998-06-03
13	PRIOR APPLICATION NUMBER: 60/088021
14	PRIOR FILING DATE: 1998-06-04
15	PRIOR APPLICATION NUMBER: 60/088025
16	PRIOR FILING DATE: 1998-06-04
17	PRIOR APPLICATION NUMBER: 60/088026
18	PRIOR FILING DATE: 1998-06-04
19	PRIOR APPLICATION NUMBER: 60/088028
20	PRIOR FILING DATE: 1998-06-04
21	PRIOR APPLICATION NUMBER: 60/088029
22	PRIOR FILING DATE: 1998-06-04
23	PRIOR APPLICATION NUMBER: 60/088030
24	PRIOR FILING DATE: 1998-06-04
25	PRIOR APPLICATION NUMBER: 60/088033
26	PRIOR FILING DATE: 1998-06-04
27	PRIOR APPLICATION NUMBER: 60/088326
28	PRIOR FILING DATE: 1998-06-04
29	PRIOR APPLICATION NUMBER: 60/088167
30	PRIOR FILING DATE: 1998-06-05
31	PRIOR APPLICATION NUMBER: 60/088202
32	PRIOR FILING DATE: 1998-06-05
33	PRIOR APPLICATION NUMBER: 60/088212
34	PRIOR FILING DATE: 1998-06-05
35	PRIOR APPLICATION NUMBER: 60/088217
36	PRIOR FILING DATE: 1998-06-05
37	PRIOR APPLICATION NUMBER: 60/088555
38	PRIOR FILING DATE: 1998-06-09
39	PRIOR APPLICATION NUMBER: 60/088734
40	PRIOR FILING DATE: 1998-06-10
41	PRIOR APPLICATION NUMBER: 60/088738
42	PRIOR FILING DATE: 1998-06-10
43	PRIOR APPLICATION NUMBER: 60/088742
44	PRIOR FILING DATE: 1998-06-10
45	PRIOR APPLICATION NUMBER: 60/088910
46	PRIOR FILING DATE: 1998-06-10
47	PRIOR APPLICATION NUMBER: 60/088824
48	PRIOR FILING DATE: 1998-06-10
49	PRIOR APPLICATION NUMBER: 60/088826
50	PRIOR FILING DATE: 1998-06-10
51	PRIOR APPLICATION NUMBER: 60/088958
52	PRIOR FILING DATE: 1998-06-11
53	PRIOR APPLICATION NUMBER: 60/088861
54	PRIOR FILING DATE: 1998-06-11
55	PRIOR APPLICATION NUMBER: 60/088876
56	PRIOR FILING DATE: 1998-06-11
57	PRIOR APPLICATION NUMBER: 60/089514
58	PRIOR FILING DATE: 1998-06-16
59	PRIOR APPLICATION NUMBER: 60/089532
60	PRIOR FILING DATE: 1998-06-17
61	PRIOR APPLICATION NUMBER: 60/089538
62	PRIOR FILING DATE: 1998-06-17
63	PRIOR APPLICATION NUMBER: 60/089598
64	PRIOR FILING DATE: 1998-06-17
65	PRIOR APPLICATION NUMBER: 60/089599
66	PRIOR FILING DATE: 1998-06-17
67	PRIOR APPLICATION NUMBER: 60/089600

```

; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          60.7%; Score 37; DB 9; Length 461;
Best Local Similarity 80.0%; Pred.No. 2.8e-02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AOEKVLQKLG 11
Db 393 AREKVLQTLG 402

RESULT 29
US-09-993-604-301
; Sequence 301, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/043787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088325
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
```

PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11  
|:|||||  
Db 393 AREKVLQTLG 402

## RESULT 30

US-09-990-456-301  
; Sequence 301, Application US/09990456  
; Patent No. US20020137890A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gottlieb, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavins, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PLC22  
; CURRENT APPLICATION NUMBER: US/09/990,456  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088025  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088555  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952

;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKUG 11

DB 393 AREKVLQTLG 402

---

; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 AOEKVLQKLG 11  
 Db 393 AREKVLQTLG 402

RESULT 32  
 US-09-992-598-301  
 ; Sequence 301, Application US/09992598  
 ; Patent No. US20020160384A1  
 ; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2730PLC20  
 ; CURRENT APPLICATION NUMBER: US/09/992,598  
 ; PRIOR FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/049787  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/065186  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066770  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/075945  
 ; PRIOR FILING DATE: 1998-02-25  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/087106  
 ; PRIOR FILING DATE: 1998-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087609  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087759  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087827  
 ; PRIOR FILING DATE: 1998-06-03  
 ; PRIOR APPLICATION NUMBER: 60/088021  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088025  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088026  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088028  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088029  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088030  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088033  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088326  
 ; PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431

PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 2 ABEKVLQKLG 11  
Db 393 ABEKVLQTLG 402

RESULT 33  
US-09-989-293A-301  
Sequence 301, Application US/09989293A  
Patent No. US20020177164A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan I.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC66  
CURRENT APPLICATION NUMBER: US/09/989,293A  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217

PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24



```

; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```

```

Query Match      60.7%; Score 37; DB 9; Length 461;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 AOEKVLQKLG 11
      |:|||||
Db      393 AREKVLQTLG 402

```

## RESULT 34

```

US-09-989-735-301
; Sequence 301, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.

```

```

; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C61
; CURRENT APPLICATION NUMBER: US/09/989,735
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10

```

; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542

; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11  
:|||||  
DB 393 AREKVLQTLG 402

RESULT 35  
US-09-990-444-301  
; Sequence 301, Application US/09990444  
; Publication No. US20020193300A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1C19  
CURRENT APPLICATION NUMBER: US/09/990,444  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;

Best Local Similarity 80.0%; Pred. No. 2.8e+02;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AQEKVLQKLG 11  
Db 393 AREKVLQTLG 402

## RESULT 36

US-09-991-181-301  
; Sequence 301, Application US/09991181  
; Publication No. US20020197615A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C53  
; CURRENT APPLICATION NUMBER: US/09/991,181

; CURRENT FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087105  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088825  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696

; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 9; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AOEKVLQKLG 11  
 :|||:|  
 Db 393 AREKVLQTLG 402

RESULT 37

US-09-989-730-301  
 ; Sequence 301, Application US/09989730  
 ; Publication No. US20020197674A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas P.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2730PIC69  
 ; CURRENT FILING DATE: 2001-11-20  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR FILING DATE: 1997-11-12

1	PRIOR APPLICATION NUMBER: 60/065311
2	PRIOR FILING DATE: 1997-11-13
3	PRIOR APPLICATION NUMBER: 60/066770
4	PRIOR FILING DATE: 1997-11-24
5	PRIOR APPLICATION NUMBER: 60/075945
6	PRIOR FILING DATE: 1998-02-25
7	PRIOR APPLICATION NUMBER: 60/078910
8	PRIOR FILING DATE: 1998-03-20
9	PRIOR APPLICATION NUMBER: 60/083322
10	PRIOR FILING DATE: 1998-04-28
11	PRIOR APPLICATION NUMBER: 60/084600
12	PRIOR FILING DATE: 1998-05-07
13	PRIOR APPLICATION NUMBER: 60/087106
14	PRIOR FILING DATE: 1998-05-28
15	PRIOR APPLICATION NUMBER: 60/087607
16	PRIOR FILING DATE: 1998-06-02
17	PRIOR APPLICATION NUMBER: 60/087609
18	PRIOR FILING DATE: 1998-06-02
19	PRIOR APPLICATION NUMBER: 60/087759
20	PRIOR FILING DATE: 1998-06-02
21	PRIOR APPLICATION NUMBER: 60/087827
22	PRIOR FILING DATE: 1998-06-03
23	PRIOR APPLICATION NUMBER: 60/088021
24	PRIOR FILING DATE: 1998-06-04
25	PRIOR APPLICATION NUMBER: 60/088025
26	PRIOR FILING DATE: 1998-06-04
27	PRIOR APPLICATION NUMBER: 60/088026
28	PRIOR FILING DATE: 1998-06-04
29	PRIOR APPLICATION NUMBER: 60/088028
30	PRIOR FILING DATE: 1998-06-04
31	PRIOR APPLICATION NUMBER: 60/088029
32	PRIOR FILING DATE: 1998-06-04
33	PRIOR APPLICATION NUMBER: 60/088030
34	PRIOR FILING DATE: 1998-06-04
35	PRIOR APPLICATION NUMBER: 60/088033
36	PRIOR FILING DATE: 1998-06-04
37	PRIOR APPLICATION NUMBER: 60/088326
38	PRIOR FILING DATE: 1998-06-04
39	PRIOR APPLICATION NUMBER: 60/088167
40	PRIOR FILING DATE: 1998-06-05
41	PRIOR APPLICATION NUMBER: 60/088202
42	PRIOR FILING DATE: 1998-06-05
43	PRIOR APPLICATION NUMBER: 60/088212
44	PRIOR FILING DATE: 1998-06-05
45	PRIOR APPLICATION NUMBER: 60/088217
46	PRIOR FILING DATE: 1998-06-05
47	PRIOR APPLICATION NUMBER: 60/088655
48	PRIOR FILING DATE: 1998-06-09
49	PRIOR APPLICATION NUMBER: 60/088734
50	PRIOR FILING DATE: 1998-06-10
51	PRIOR APPLICATION NUMBER: 60/088738
52	PRIOR FILING DATE: 1998-06-10
53	PRIOR APPLICATION NUMBER: 60/088742
54	PRIOR FILING DATE: 1998-06-10
55	PRIOR APPLICATION NUMBER: 60/088810
56	PRIOR FILING DATE: 1998-06-10
57	PRIOR APPLICATION NUMBER: 60/088824
58	PRIOR FILING DATE: 1998-06-10
59	PRIOR APPLICATION NUMBER: 60/088826
60	PRIOR FILING DATE: 1998-06-10
61	PRIOR APPLICATION NUMBER: 60/088858
62	PRIOR FILING DATE: 1998-06-11
63	PRIOR APPLICATION NUMBER: 60/088861
64	PRIOR FILING DATE: 1998-06-11
65	PRIOR APPLICATION NUMBER: 60/088876
66	PRIOR FILING DATE: 1998-06-11
67	PRIOR APPLICATION NUMBER: 60/089105
68	PRIOR FILING DATE: 1998-06-12
69	PRIOR APPLICATION NUMBER: 60/089440
70	PRIOR FILING DATE: 1998-06-16
71	PRIOR APPLICATION NUMBER: 60/089512
72	PRIOR FILING DATE: 1998-06-16
73	PRIOR APPLICATION NUMBER: 60/089514

```

; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          60.7%; Score 37; DB 9; Length 461;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 AQEVLQKLG 11
Db      393 AREVLQTLG 402

RESULT 38
US-09-990-436-301
; Sequence 301, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tuma, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC14
; CURRENT APPLICATION NUMBER: US/09/990,436
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17

```

; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089507  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626

; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7% Score 37; DB 9; Length 461;  
 Best local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 AQEYVQLKLG 11  
 : : : : :  
 DB 393 AREKVLQTLG 402

RESULT 39  
 US-09-993-687-301  
 ; Sequence 301, Application US/09993687  
 ; Publication No. US20020198149A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2730PIC11  
 ; CURRENT APPLICATION NUMBER: US/09/993.687  
 ; CURRENT FILING DATE: 2002-11-14  
 ; PRIOR APPLICATION NUMBER: 60/049787  
 ; PRIOR FILING DATE: 1997-06-15  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/065186  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066770  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/075945  
 ; PRIOR FILING DATE: 1998-02-25  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/087106  
 ; PRIOR FILING DATE: 1998-05-28



1 PRIOR APPLICATION NUMBER: 60/087607  
2 PRIOR FILING DATE: 1998-06-02  
3 PRIOR APPLICATION NUMBER: 60/087609  
4 PRIOR FILING DATE: 1998-06-02  
5 PRIOR APPLICATION NUMBER: 60/087759  
6 PRIOR FILING DATE: 1998-06-02  
7 PRIOR APPLICATION NUMBER: 60/087827  
8 PRIOR FILING DATE: 1998-06-03  
9 PRIOR APPLICATION NUMBER: 60/088021  
10 PRIOR FILING DATE: 1998-06-04  
11 PRIOR APPLICATION NUMBER: 60/088025  
12 PRIOR FILING DATE: 1998-06-04  
13 PRIOR APPLICATION NUMBER: 60/088026  
14 PRIOR FILING DATE: 1998-06-04  
15 PRIOR APPLICATION NUMBER: 60/088028  
16 PRIOR FILING DATE: 1998-06-04  
17 PRIOR APPLICATION NUMBER: 60/088029  
18 PRIOR FILING DATE: 1998-06-04  
19 PRIOR APPLICATION NUMBER: 60/088030  
20 PRIOR FILING DATE: 1998-06-04  
21 PRIOR APPLICATION NUMBER: 60/088033  
22 PRIOR FILING DATE: 1998-06-04  
23 PRIOR APPLICATION NUMBER: 60/088326  
24 PRIOR FILING DATE: 1998-06-04  
25 PRIOR APPLICATION NUMBER: 60/088167  
26 PRIOR FILING DATE: 1998-06-05  
27 PRIOR APPLICATION NUMBER: 60/088202  
28 PRIOR FILING DATE: 1998-06-05  
29 PRIOR APPLICATION NUMBER: 60/088212  
30 PRIOR FILING DATE: 1998-06-05  
31 PRIOR APPLICATION NUMBER: 60/088217  
32 PRIOR FILING DATE: 1998-06-05  
33 PRIOR APPLICATION NUMBER: 60/088655  
34 PRIOR FILING DATE: 1998-06-09  
35 PRIOR APPLICATION NUMBER: 60/088734  
36 PRIOR FILING DATE: 1998-06-10  
37 PRIOR APPLICATION NUMBER: 60/088738  
38 PRIOR FILING DATE: 1998-06-10  
39 PRIOR APPLICATION NUMBER: 60/088742  
40 PRIOR FILING DATE: 1998-06-10  
41 PRIOR APPLICATION NUMBER: 60/088810  
42 PRIOR FILING DATE: 1998-06-10  
43 PRIOR APPLICATION NUMBER: 60/088824  
44 PRIOR FILING DATE: 1998-06-10  
45 PRIOR APPLICATION NUMBER: 60/088826  
46 PRIOR FILING DATE: 1998-06-10  
47 PRIOR APPLICATION NUMBER: 60/088858  
48 PRIOR FILING DATE: 1998-06-11  
49 PRIOR APPLICATION NUMBER: 60/088861  
50 PRIOR FILING DATE: 1998-06-11  
51 PRIOR APPLICATION NUMBER: 60/088876  
52 PRIOR FILING DATE: 1998-06-11  
53 PRIOR APPLICATION NUMBER: 60/089105  
54 PRIOR FILING DATE: 1998-06-12  
55 PRIOR APPLICATION NUMBER: 60/089440  
56 PRIOR FILING DATE: 1998-06-16  
57 PRIOR APPLICATION NUMBER: 60/089512  
58 PRIOR FILING DATE: 1998-06-16  
59 PRIOR APPLICATION NUMBER: 60/089514  
60 PRIOR FILING DATE: 1998-06-16  
61 PRIOR APPLICATION NUMBER: 60/089532  
62 PRIOR FILING DATE: 1998-06-17  
63 PRIOR APPLICATION NUMBER: 60/089538  
64 PRIOR FILING DATE: 1998-06-17  
65 PRIOR APPLICATION NUMBER: 60/089598  
66 PRIOR FILING DATE: 1998-06-17  
67 PRIOR APPLICATION NUMBER: 60/089599  
68 PRIOR FILING DATE: 1998-06-17  
69 PRIOR APPLICATION NUMBER: 60/089600  
70 PRIOR FILING DATE: 1998-06-17  
71 PRIOR APPLICATION NUMBER: 60/089653  
72 PRIOR FILING DATE: 1998-06-17  
73 PRIOR APPLICATION NUMBER: 60/089801  
74 PRIOR FILING DATE: 1998-06-18  
75 PRIOR APPLICATION NUMBER: 60/089907  
76 PRIOR FILING DATE: 1998-06-18  
77 PRIOR APPLICATION NUMBER: 60/089908  
78 PRIOR FILING DATE: 1998-06-18  
79 PRIOR APPLICATION NUMBER: 60/089947  
80 PRIOR FILING DATE: 1998-06-19  
81 PRIOR APPLICATION NUMBER: 60/089948  
82 PRIOR FILING DATE: 1998-06-19  
83 PRIOR APPLICATION NUMBER: 60/089952  
84 PRIOR FILING DATE: 1998-06-19  
85 PRIOR APPLICATION NUMBER: 60/090246  
86 PRIOR FILING DATE: 1998-06-22  
87 PRIOR APPLICATION NUMBER: 60/090252  
88 PRIOR FILING DATE: 1998-06-22  
89 PRIOR APPLICATION NUMBER: 60/090254  
90 PRIOR FILING DATE: 1998-06-22  
91 PRIOR APPLICATION NUMBER: 60/090349  
92 PRIOR FILING DATE: 1998-06-23  
93 PRIOR APPLICATION NUMBER: 60/090355  
94 PRIOR FILING DATE: 1998-06-23  
95 PRIOR APPLICATION NUMBER: 60/090429  
96 PRIOR FILING DATE: 1998-06-24  
97 PRIOR APPLICATION NUMBER: 60/090431  
98 PRIOR FILING DATE: 1998-06-24  
99 PRIOR APPLICATION NUMBER: 60/090435  
100 PRIOR FILING DATE: 1998-06-24  
101 PRIOR APPLICATION NUMBER: 60/090444  
102 PRIOR FILING DATE: 1998-06-24  
103 PRIOR APPLICATION NUMBER: 60/090445  
104 PRIOR FILING DATE: 1998-06-24  
105 PRIOR APPLICATION NUMBER: 60/090472  
106 PRIOR FILING DATE: 1998-06-24  
107 PRIOR APPLICATION NUMBER: 60/090535  
108 PRIOR FILING DATE: 1998-06-24  
109 PRIOR APPLICATION NUMBER: 60/090540  
110 PRIOR FILING DATE: 1998-06-24  
111 PRIOR APPLICATION NUMBER: 60/090542  
112 PRIOR FILING DATE: 1998-06-24  
113 PRIOR APPLICATION NUMBER: 60/090557  
114 PRIOR FILING DATE: 1998-06-24  
115 PRIOR APPLICATION NUMBER: 60/090676  
116 PRIOR FILING DATE: 1998-06-25  
117 PRIOR APPLICATION NUMBER: 60/090678  
118 PRIOR FILING DATE: 1998-06-25  
119 PRIOR APPLICATION NUMBER: 60/090690  
120 PRIOR FILING DATE: 1998-06-25  
121 PRIOR APPLICATION NUMBER: 60/090694  
122 PRIOR FILING DATE: 1998-06-25  
123 PRIOR APPLICATION NUMBER: 60/090695  
124 PRIOR FILING DATE: 1998-06-25  
125 PRIOR APPLICATION NUMBER: 60/090696  
126 PRIOR FILING DATE: 1998-06-25  
127 PRIOR APPLICATION NUMBER: 60/090862  
128 PRIOR FILING DATE: 1998-06-26  
129 PRIOR APPLICATION NUMBER: 60/090863  
130 PRIOR FILING DATE: 1998-06-26  
131 PRIOR APPLICATION NUMBER: 60/091360  
132 PRIOR FILING DATE: 1998-07-01  
133 PRIOR APPLICATION NUMBER: 60/091478  
134 PRIOR FILING DATE: 1998-07-02  
135 PRIOR APPLICATION NUMBER: 60/091544  
136 PRIOR FILING DATE: 1998-07-01  
137 PRIOR APPLICATION NUMBER: 60/091519  
138 PRIOR FILING DATE: 1998-07-02  
139 PRIOR APPLICATION NUMBER: 60/091626  
140 PRIOR FILING DATE: 1998-07-02  
141 PRIOR APPLICATION NUMBER: 60/091633  
142 PRIOR FILING DATE: 1998-07-02  
143 PRIOR APPLICATION NUMBER: 60/091978  
144 PRIOR FILING DATE: 1998-07-07  
145 PRIOR APPLICATION NUMBER: 60/091982  
146 PRIOR FILING DATE: 1998-07-07

```
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      60.7%; Score 37; DB 9; Length 461;
Best Local Similarity 80.0%; Pred. NO. 2.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 AOEKVLQKLG 11
      1:|||||
Db      393 AREKVLQILG 402

RESULT 40
US-09-989-734-301
; Sequence 301, Application US/09989734
; Publication No. US20030003531A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Wickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC64
; CURRENT APPLICATION NUMBER: US/09/989,734
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827

; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
```

; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 10; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AOEKVLOKLG 11  
 Db 393 AREKVLOTLG 402  
 RESULT 41  
 US-09-997-653-301  
 ; Sequence 301, Application US/09997653  
 ; Publication No. US20030008297A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Deenoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Kijavini, Ivar J.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tamas, Daniel  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2730P1C38  
 ; CURRENT APPLICATION NUMBER: US/09/997,653  
 ; CURRENT FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: 60/049787  
 ; PRIOR FILING DATE: 1997-06-16  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/065186  
 ; PRIOR FILING DATE: 1997-11-12  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066770  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/075945  
 ; PRIOR FILING DATE: 1998-02-25  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/087106  
 ; PRIOR FILING DATE: 1998-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087607  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087609  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087759  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087827  
 ; PRIOR FILING DATE: 1998-06-03  
 ; PRIOR APPLICATION NUMBER: 60/088021  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088025  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088026  
 ; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088028  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088029  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088030  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088033  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088326  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088167  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088202  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088212  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088217  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088655  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: 60/088734  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088738  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088742  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088810  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088858  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252

; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 10; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Caps 0;

QY 2 AOEKVLQKLG 11  
 Db 393 ABEKVLQKLG 402

RESULT 42  
 US-09-993-667-301

; Sequence 301, Application US/05993667  
; Publication No. US2003022187A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC4  
; CURRENT APPLICATION NUMBER: US/09/993,667  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23

APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas P.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C44  
CURRENT APPLICATION NUMBER: US/09/997,428  
CURRENT FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05

Query Match 60.7%; Score 37; DB 10; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AOEKVLQKLG 11

Db 393 AREKVLQTLG 402

RESULT 43

US-09-997-428-301  
Sequence 301, Application US/09997428  
Publication No. US20030027162A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnucy, Luc

; PRIOR APPLICATION NUMBER: 60/088212  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088217  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088655  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: 60/088734  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088738  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088742  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088810  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088858  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444

; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 10; Length 461;  
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AOEKVLQKLG 11  
 | : | | | | | | | |  
 Db 393 AREKVLQTLG 402

RESULT 44

US-09-997-666-301  
 ; Sequence 301, Application US/09997666  
 ; Publication No. US20030027163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi J.  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijaviri, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C42  
CURRENT APPLICATION NUMBER: US/09/997,666  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24



; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 10; Length 461;

Best Local Similarity 80.0%; Pred. No. 2.8e+02;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 AOEKVLQKLG 11

Db 393 AREKVLQTLG 402

#### RESULT 45

US-03-990-438-301  
; Sequence 301, Application US/09990438  
; Publication No. US2003002754A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Geritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC3  
; CURRENT APPLICATION NUMBER: US/09/990,438  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088858  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089440  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089532  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089599  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089600  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089801  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089907  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 ; PRIOR FILING DATE: 1998-06-18  
 ; PRIOR APPLICATION NUMBER: 60/089947  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089948  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/089952  
 ; PRIOR FILING DATE: 1998-06-19  
 ; PRIOR APPLICATION NUMBER: 60/090246  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090252  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090254  
 ; PRIOR FILING DATE: 1998-06-22  
 ; PRIOR APPLICATION NUMBER: 60/090349  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090355  
 ; PRIOR FILING DATE: 1998-06-23  
 ; PRIOR APPLICATION NUMBER: 60/090429  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090431  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090435  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090444  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090445  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090472  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090535  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090540  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090542  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090557  
 ; PRIOR FILING DATE: 1998-06-24  
 ; PRIOR APPLICATION NUMBER: 60/090676

; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090678  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090690  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090694  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090695  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090696  
 ; PRIOR FILING DATE: 1998-06-25  
 ; PRIOR APPLICATION NUMBER: 60/090862  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/090863  
 ; PRIOR FILING DATE: 1998-06-26  
 ; PRIOR APPLICATION NUMBER: 60/091360  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091478  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091544  
 ; PRIOR FILING DATE: 1998-07-01  
 ; PRIOR APPLICATION NUMBER: 60/091519  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091626  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091633  
 ; PRIOR FILING DATE: 1998-07-02  
 ; PRIOR APPLICATION NUMBER: 60/091978  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/091982  
 ; PRIOR FILING DATE: 1998-07-07  
 ; PRIOR APPLICATION NUMBER: 60/092182  
 ; PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 10; Length 461;

Best Local Similarity 80.0%; Pred. No. 2.8e+02;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 AOEKVLQKLG 11

Db 393 AREKVLQTLG 402

RESULT 46

US-08-990-562-301

; Sequence 301, Application US/09990562

; Publication No. US20030027985A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

1 TITLE OF INVENTION: Acids Encoding the Same  
2 FILE REFERENCE: P2730PLC18  
3 CURRENT APPLICATION NUMBER: US/09/390,562  
4 CURRENT FILING DATE: 2001-11-14  
5 PRIOR APPLICATION NUMBER: 60/049787  
6 PRIOR FILING DATE: 1997-06-16  
7 PRIOR APPLICATION NUMBER: 60/062250  
8 PRIOR FILING DATE: 1997-10-17  
9 PRIOR APPLICATION NUMBER: 60/065186  
10 PRIOR FILING DATE: 1997-11-12  
11 PRIOR APPLICATION NUMBER: 60/065311  
12 PRIOR FILING DATE: 1997-11-13  
13 PRIOR APPLICATION NUMBER: 60/066770  
14 PRIOR FILING DATE: 1997-11-24  
15 PRIOR APPLICATION NUMBER: 60/075945  
16 PRIOR FILING DATE: 1998-02-25  
17 PRIOR APPLICATION NUMBER: 60/078910  
18 PRIOR FILING DATE: 1998-03-20  
19 PRIOR APPLICATION NUMBER: 60/083322  
20 PRIOR FILING DATE: 1998-04-28  
21 PRIOR APPLICATION NUMBER: 60/084600  
22 PRIOR FILING DATE: 1998-05-07  
23 PRIOR APPLICATION NUMBER: 60/087106  
24 PRIOR FILING DATE: 1998-05-28  
25 PRIOR APPLICATION NUMBER: 60/087607  
26 PRIOR FILING DATE: 1998-06-02  
27 PRIOR APPLICATION NUMBER: 60/087609  
28 PRIOR FILING DATE: 1998-06-02  
29 PRIOR APPLICATION NUMBER: 60/087759  
30 PRIOR FILING DATE: 1998-06-02  
31 PRIOR APPLICATION NUMBER: 60/087827  
32 PRIOR FILING DATE: 1998-06-03  
33 PRIOR APPLICATION NUMBER: 60/088021  
34 PRIOR FILING DATE: 1998-06-04  
35 PRIOR APPLICATION NUMBER: 60/088025  
36 PRIOR FILING DATE: 1998-06-04  
37 PRIOR APPLICATION NUMBER: 60/088026  
38 PRIOR FILING DATE: 1998-06-04  
39 PRIOR APPLICATION NUMBER: 60/088028  
40 PRIOR FILING DATE: 1998-06-04  
41 PRIOR APPLICATION NUMBER: 60/088029  
42 PRIOR FILING DATE: 1998-06-04  
43 PRIOR APPLICATION NUMBER: 60/088030  
44 PRIOR FILING DATE: 1998-06-04  
45 PRIOR APPLICATION NUMBER: 60/088033  
46 PRIOR FILING DATE: 1998-06-04  
47 PRIOR APPLICATION NUMBER: 60/088326  
48 PRIOR FILING DATE: 1998-06-04  
49 PRIOR APPLICATION NUMBER: 60/088167  
50 PRIOR FILING DATE: 1998-06-05  
51 PRIOR APPLICATION NUMBER: 60/088202  
52 PRIOR FILING DATE: 1998-06-05  
53 PRIOR APPLICATION NUMBER: 60/088212  
54 PRIOR FILING DATE: 1998-06-05  
55 PRIOR APPLICATION NUMBER: 60/088217  
56 PRIOR FILING DATE: 1998-06-05  
57 PRIOR APPLICATION NUMBER: 60/088655  
58 PRIOR FILING DATE: 1998-06-09  
59 PRIOR APPLICATION NUMBER: 60/088734  
60 PRIOR FILING DATE: 1998-06-10  
61 PRIOR APPLICATION NUMBER: 60/088738  
62 PRIOR FILING DATE: 1998-06-10  
63 PRIOR APPLICATION NUMBER: 60/088742  
64 PRIOR FILING DATE: 1998-06-10  
65 PRIOR APPLICATION NUMBER: 60/088810  
66 PRIOR FILING DATE: 1998-06-10  
67 PRIOR APPLICATION NUMBER: 60/088824  
68 PRIOR FILING DATE: 1998-06-10  
69 PRIOR APPLICATION NUMBER: 60/088826  
70 PRIOR FILING DATE: 1998-06-10  
71 PRIOR APPLICATION NUMBER: 60/088858  
72 PRIOR FILING DATE: 1998-06-11  
73 PRIOR APPLICATION NUMBER: 60/088861  
74 PRIOR FILING DATE: 1998-06-11  
75 PRIOR APPLICATION NUMBER: 60/088876  
76 PRIOR FILING DATE: 1998-06-25  
77 PRIOR APPLICATION NUMBER: 60/090678  
78 PRIOR FILING DATE: 1998-06-25  
79 PRIOR APPLICATION NUMBER: 60/090690  
80 PRIOR FILING DATE: 1998-06-25  
81 PRIOR APPLICATION NUMBER: 60/090694  
82 PRIOR FILING DATE: 1998-06-25  
83 PRIOR APPLICATION NUMBER: 60/088876  
84 PRIOR FILING DATE: 1998-06-25  
85 PRIOR APPLICATION NUMBER: 60/090678  
86 PRIOR FILING DATE: 1998-06-25  
87 PRIOR APPLICATION NUMBER: 60/090542  
88 PRIOR FILING DATE: 1998-06-24  
89 PRIOR APPLICATION NUMBER: 60/090557  
90 PRIOR FILING DATE: 1998-06-24  
91 PRIOR APPLICATION NUMBER: 60/090676  
92 PRIOR FILING DATE: 1998-06-25  
93 PRIOR APPLICATION NUMBER: 60/090678  
94 PRIOR FILING DATE: 1998-06-25  
95 PRIOR APPLICATION NUMBER: 60/090690  
96 PRIOR FILING DATE: 1998-06-25  
97 PRIOR APPLICATION NUMBER: 60/090694  
98 PRIOR FILING DATE: 1998-06-25  
99 PRIOR APPLICATION NUMBER: 60/088876  
100 PRIOR FILING DATE: 1998-06-11  
101 PRIOR APPLICATION NUMBER: 60/088876  
102 PRIOR FILING DATE: 1998-06-11  
103 PRIOR APPLICATION NUMBER: 60/089105  
104 PRIOR FILING DATE: 1998-06-12  
105 PRIOR APPLICATION NUMBER: 60/089440  
106 PRIOR FILING DATE: 1998-06-16  
107 PRIOR APPLICATION NUMBER: 60/089512  
108 PRIOR FILING DATE: 1998-06-16  
109 PRIOR APPLICATION NUMBER: 60/089514  
110 PRIOR FILING DATE: 1998-06-16  
111 PRIOR APPLICATION NUMBER: 60/089532  
112 PRIOR FILING DATE: 1998-06-17  
113 PRIOR APPLICATION NUMBER: 60/089538  
114 PRIOR FILING DATE: 1998-06-17  
115 PRIOR APPLICATION NUMBER: 60/089598  
116 PRIOR FILING DATE: 1998-06-17  
117 PRIOR APPLICATION NUMBER: 60/089599  
118 PRIOR FILING DATE: 1998-06-17  
119 PRIOR APPLICATION NUMBER: 60/089600  
120 PRIOR FILING DATE: 1998-06-17  
121 PRIOR APPLICATION NUMBER: 60/089653  
122 PRIOR FILING DATE: 1998-06-17  
123 PRIOR APPLICATION NUMBER: 60/089801  
124 PRIOR FILING DATE: 1998-06-18  
125 PRIOR APPLICATION NUMBER: 60/089907  
126 PRIOR FILING DATE: 1998-06-18  
127 PRIOR APPLICATION NUMBER: 60/089908  
128 PRIOR FILING DATE: 1998-06-18  
129 PRIOR APPLICATION NUMBER: 60/089947  
130 PRIOR FILING DATE: 1998-06-19  
131 PRIOR APPLICATION NUMBER: 60/089948  
132 PRIOR FILING DATE: 1998-06-19  
133 PRIOR APPLICATION NUMBER: 60/089952  
134 PRIOR FILING DATE: 1998-06-19  
135 PRIOR APPLICATION NUMBER: 60/090246  
136 PRIOR FILING DATE: 1998-06-22  
137 PRIOR APPLICATION NUMBER: 60/090252  
138 PRIOR FILING DATE: 1998-06-22  
139 PRIOR APPLICATION NUMBER: 60/090254  
140 PRIOR FILING DATE: 1998-06-22  
141 PRIOR APPLICATION NUMBER: 60/090349  
142 PRIOR FILING DATE: 1998-06-23  
143 PRIOR APPLICATION NUMBER: 60/090355  
144 PRIOR FILING DATE: 1998-06-23  
145 PRIOR APPLICATION NUMBER: 60/090429  
146 PRIOR FILING DATE: 1998-06-24  
147 PRIOR APPLICATION NUMBER: 60/090431  
148 PRIOR FILING DATE: 1998-06-24  
149 PRIOR APPLICATION NUMBER: 60/090435  
150 PRIOR FILING DATE: 1998-06-24  
151 PRIOR APPLICATION NUMBER: 60/090444  
152 PRIOR FILING DATE: 1998-06-24  
153 PRIOR APPLICATION NUMBER: 60/090445  
154 PRIOR FILING DATE: 1998-06-24  
155 PRIOR APPLICATION NUMBER: 60/090472  
156 PRIOR FILING DATE: 1998-06-24  
157 PRIOR APPLICATION NUMBER: 60/090535  
158 PRIOR FILING DATE: 1998-06-24  
159 PRIOR APPLICATION NUMBER: 60/090540  
160 PRIOR FILING DATE: 1998-06-24  
161 PRIOR APPLICATION NUMBER: 60/090542  
162 PRIOR FILING DATE: 1998-06-24  
163 PRIOR APPLICATION NUMBER: 60/090557  
164 PRIOR FILING DATE: 1998-06-24  
165 PRIOR APPLICATION NUMBER: 60/090676  
166 PRIOR FILING DATE: 1998-06-25  
167 PRIOR APPLICATION NUMBER: 60/090678  
168 PRIOR FILING DATE: 1998-06-25  
169 PRIOR APPLICATION NUMBER: 60/090690  
170 PRIOR FILING DATE: 1998-06-25  
171 PRIOR APPLICATION NUMBER: 60/090694  
172 PRIOR FILING DATE: 1998-06-25

```

; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      60.7%; Score 37; DB 10; Length 461;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AQEKVLQKLG 11
Db      393 AREKVLQTLG 402

RESULT 47
US-09-990-711-301
; Sequence 301; Application US/09990711
; Publication No. US20030302023A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC2
; CURRENT APPLICATION NUMBER: US/09/990,711
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-15

```

1 PRIOR APPLICATION NUMBER: 60/089512  
2 PRIOR FILING DATE: 1998-06-16  
3 PRIOR APPLICATION NUMBER: 60/089514  
4 PRIOR FILING DATE: 1998-06-16  
5 PRIOR APPLICATION NUMBER: 60/089532  
6 PRIOR FILING DATE: 1998-06-17  
7 PRIOR APPLICATION NUMBER: 60/089538  
8 PRIOR FILING DATE: 1998-06-17  
9 PRIOR APPLICATION NUMBER: 60/089598  
10 PRIOR FILING DATE: 1998-06-17  
11 PRIOR APPLICATION NUMBER: 60/089599  
12 PRIOR FILING DATE: 1998-06-17  
13 PRIOR APPLICATION NUMBER: 60/089600  
14 PRIOR FILING DATE: 1998-06-17  
15 PRIOR APPLICATION NUMBER: 60/089653  
16 PRIOR FILING DATE: 1998-06-17  
17 PRIOR APPLICATION NUMBER: 60/089801  
18 PRIOR FILING DATE: 1998-06-18  
19 PRIOR APPLICATION NUMBER: 60/089907  
20 PRIOR FILING DATE: 1998-06-18  
21 PRIOR APPLICATION NUMBER: 60/089908  
22 PRIOR FILING DATE: 1998-06-18  
23 PRIOR APPLICATION NUMBER: 60/089947  
24 PRIOR FILING DATE: 1998-06-19  
25 PRIOR APPLICATION NUMBER: 60/089948  
26 PRIOR FILING DATE: 1998-06-19  
27 PRIOR APPLICATION NUMBER: 60/089952  
28 PRIOR FILING DATE: 1998-06-19  
29 PRIOR APPLICATION NUMBER: 60/090246  
30 PRIOR FILING DATE: 1998-06-22  
31 PRIOR APPLICATION NUMBER: 60/090252  
32 PRIOR FILING DATE: 1998-06-22  
33 PRIOR APPLICATION NUMBER: 60/090254  
34 PRIOR FILING DATE: 1998-06-22  
35 PRIOR APPLICATION NUMBER: 60/090349  
36 PRIOR FILING DATE: 1998-06-23  
37 PRIOR APPLICATION NUMBER: 60/090355  
38 PRIOR FILING DATE: 1998-06-23  
39 PRIOR APPLICATION NUMBER: 60/090429  
40 PRIOR FILING DATE: 1998-06-24  
41 PRIOR APPLICATION NUMBER: 60/090431  
42 PRIOR FILING DATE: 1998-06-24  
43 PRIOR APPLICATION NUMBER: 60/090435  
44 PRIOR FILING DATE: 1998-06-24  
45 PRIOR APPLICATION NUMBER: 60/090444  
46 PRIOR FILING DATE: 1998-06-24  
47 PRIOR APPLICATION NUMBER: 60/090445  
48 PRIOR FILING DATE: 1998-06-24  
49 PRIOR APPLICATION NUMBER: 60/090472  
50 PRIOR FILING DATE: 1998-06-24  
51 PRIOR APPLICATION NUMBER: 60/090535  
52 PRIOR FILING DATE: 1998-06-24  
53 PRIOR APPLICATION NUMBER: 60/090540  
54 PRIOR FILING DATE: 1998-06-24  
55 PRIOR APPLICATION NUMBER: 60/090542  
56 PRIOR FILING DATE: 1998-06-24  
57 PRIOR APPLICATION NUMBER: 60/090557  
58 PRIOR FILING DATE: 1998-06-24  
59 PRIOR APPLICATION NUMBER: 60/090676  
60 PRIOR FILING DATE: 1998-06-25  
61 PRIOR APPLICATION NUMBER: 60/090678  
62 PRIOR FILING DATE: 1998-06-25  
63 PRIOR APPLICATION NUMBER: 60/090690  
64 PRIOR FILING DATE: 1998-06-25  
65 PRIOR APPLICATION NUMBER: 60/090694  
66 PRIOR FILING DATE: 1998-06-25  
67 PRIOR APPLICATION NUMBER: 60/090695  
68 PRIOR FILING DATE: 1998-06-25  
69 PRIOR APPLICATION NUMBER: 60/090696  
70 PRIOR FILING DATE: 1998-06-25  
71 PRIOR APPLICATION NUMBER: 60/090862  
72 PRIOR FILING DATE: 1998-06-26  
73 PRIOR APPLICATION NUMBER: 60/090863

1 PRIOR FILING DATE: 1998-06-26  
2 PRIOR APPLICATION NUMBER: 60/091360  
3 PRIOR FILING DATE: 1998-07-01  
4 PRIOR APPLICATION NUMBER: 60/091478  
5 PRIOR FILING DATE: 1998-07-02  
6 PRIOR APPLICATION NUMBER: 60/091544  
7 PRIOR FILING DATE: 1998-07-01  
8 PRIOR APPLICATION NUMBER: 60/091519  
9 PRIOR FILING DATE: 1998-07-02  
10 PRIOR APPLICATION NUMBER: 60/091626  
11 PRIOR FILING DATE: 1998-07-02  
12 PRIOR APPLICATION NUMBER: 60/091633  
13 PRIOR FILING DATE: 1998-07-02  
14 PRIOR APPLICATION NUMBER: 60/091978  
15 PRIOR FILING DATE: 1998-07-07  
16 PRIOR APPLICATION NUMBER: 60/091982  
17 PRIOR FILING DATE: 1998-07-07  
18 PRIOR APPLICATION NUMBER: 60/092182  
19 PRIOR FILING DATE: 1998-07-09

Query Match 60.7%; Score 37; DB 10; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AOEKVLQKLG 11  
|:|||||  
DB 393 AREKVLQTLG 402

## RESULT 48

US-09-989-726-301  
; Sequence 301, Application US/09989726  
; Publication No. US2003004073A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas P.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730F1C60  
; CURRENT APPLICATION NUMBER: US/09/989,726  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24



```
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          60.7%; Score 37; DB 10; Length 461;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  2 AOEKVLQKLG 11
Db  393 AREKVLQTLG 402

RESULT 49
US-09-998-156-301
; Sequence 301, Application US/09998156
; Publication No. US20030044806A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P27301C28
; CURRENT APPLICATION NUMBER: US/09/998,156
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/043787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
```

; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978

; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 60.7% Score 37; DB 10; Length 461;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLG 11  
|:|||||  
Db 393 AREKVLQTLG 402

## RESULT 50

US-09-990-437-301  
; Sequence 301, Application US/09990437  
; Publication No. US20030045463A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C49  
; CURRENT APPLICATION NUMBER: US/09/990,437  
; CURRENT FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02





Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AOEKVLQKLG 11  
|:|||||  
Db 393 AREKVLQTLG 402

Search completed: March 4, 2004, 17:55:15  
Job time : 27.0845 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:24:55 ; Search time 9.29045 Seconds  
(without alignments)  
1228.072 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_1\_221  
Perfect score: 1125  
Sequence: 1 MAEGKAGGAGLFAKQVQK.....IGCVTIFQNISLRDVFYR 221

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgm2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/2/iaa/6C\_COMB.pep.\*  
6: /cgm2\_6/ptodata/2/iaa/6D\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	689	61.2	434	US-08-630-915A-22	Sequence 22, Appl
2	671.5	59.7	451	US-08-435-454-4	Sequence 4, Appli
3	671.5	59.7	451	US-08-652-972A-4	Sequence 4, Appli
4	671.5	59.7	451	US-08-919-145-6	Sequence 6, Appli
5	671.5	59.7	451	US-08-870-126-4	Sequence 4, Appli
6	671.5	59.7	451	US-09-344-889-6	Sequence 6, Appli
7	671.5	59.7	451	US-09-445-247-4	Sequence 4, Appli
8	671.5	59.7	451	PCT-US96-06231A-4	Sequence 4, Appli
9	559.5	49.7	404	US-08-630-915A-24	Sequence 24, Appl
10	113	10.0	1972	US-08-875-435B-3	Sequence 3, Appli
11	111.5	9.9	2954	US-09-150-867-1	Sequence 1, Appli
12	110	9.8	816	US-08-533-306A-6	Sequence 6, Appli
13	110	9.8	816	US-08-742-923A-6	Sequence 6, Appli
14	110	9.8	885	US-08-533-306A-4	Sequence 6, Appli
15	110	9.8	885	US-08-742-923A-4	Sequence 4, Appli
16	108	9.6	1972	US-08-875-435B-4	Sequence 4, Appli
17	107	9.5	975	US-09-914-259-19	Sequence 19, Appl
18	107	9.5	8991	US-08-714-741-32	Sequence 32, Appl
19	106.5	9.5	2285	US-09-308-375-2	Sequence 2, Appli
20	104	9.2	1581	US-09-866-108A-15754	Sequence 15754, A
21	103.5	9.2	2551	US-09-328-352-8156	Sequence 8156, Ap
22	103.5	9.2	818	US-09-134-000C-6355	Sequence 6355, Ap
23	103.5	9.2	906	US-09-107-532A-6450	Sequence 6450, Ap
24	103.5	9.2	1010	US-09-134-001C-5178	Sequence 5178, Ap
25	103	9.2	967	US-09-914-259-21	Sequence 21, Appl
26	103	9.2	2662	US-09-595-684B-31	Sequence 31, Appl
27	103	9.2	3878	US-09-914-259-11	Sequence 11, Appl

Sequence 558, App  
Sequence 4820, Ap  
Sequence 5466, Ap  
Sequence 5, Appli  
Sequence 15753, A  
Sequence 5352, Ap  
Sequence 220, App  
Sequence 5067, Ap  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 11, Appl  
Sequence 5178, Ap  
Sequence 53, Appli  
Sequence 2, Appli  
Sequence 5756, Ap

## ALIGNMENTS

RESULT 1  
US-08-630-915A-22  
; Sequence 22, Application US/08630915A  
; Patent No. 6309820  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: HOFFMAN, No. 6309820h  
; APPLICANT: KAY, Brian K.  
; APPLICANT: FOWLES, Dana M.  
; APPLICANT: MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; TITLE OF INVENTION: USING SAME  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/630,915A  
; FILING DATE: 03-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirostock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-174  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 434 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-915A-22

Query Match 61.2%, Score 689, DB 4, Length 434;  
Best Local Similarity 61.9%, Pred. No. 1.3e-54;  
Matches 138; Conservative 30; Mismatches 53; Indels 2; Gaps 2;

QY 1 MAE-GKAGGAGLFAKQVQKFSRAQEKVLQKLGAVETKDERPEQASNFYQQAEGHK 59  
 Db 1 MAEMSGKGTAGKIASNVQKLTQAEKVLQKLGKADTKDEQEQCVQNFKNKLTGTR 60  
 QY 60 LYKDLNFLSAVKVHSHSKRVSSETLQEIYSSEWMDGHEELKAIWVNDLLWEDYEELAD 119  
 Db 61 LQKDLRTYLSAVKAMHEASKLSECLQEVYEPWFGDEANKIAENNDLLWMDYHQKLV 120  
 QY 120 QAVRTWEIYVAQSEIKERIAKRGKLVYDSARHHLAVQNA-KDEAKTAKAEFEFNK 178  
 Db 121 QALLTMDTYLGGQFPDIKSRKAGKLVYDSARHHLAVQNA-KDEAKTAKAEFEFNK 180  
 QY 179 AQTVPEDLNQELLELPILYNSRICGYVTIFQNISNLRDVYR 221  
 Db 181 AQVFEEMVNDLQBELPWSNRVGVYNTFQSIAGLEENFK 223

## RESULT 2

US-08-435-454-4  
 ; Sequence 4, Application US/08435454  
 ; Patent No. 5605830  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prndergast, George C.  
 ; APPLICANT: Sakamuro, Daitoku  
 ; TITLE OF INVENTION: Murine and Human C-Myc Interacting  
 ; TITLE OF INVENTION: Protein and Uses Therefor  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P O Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/435,454  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bak, Mary E.  
 ; REGISTRATION NUMBER: 31,215  
 ; REFERENCE/DOCKET NUMBER: WST60USA  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-435-454-4

Query Match 59.7%; Score 671.5; DB 1; Length 451;  
 Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
 Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;  
 QY 10 AGLEAKQVQKFSRAQEKVLQKLGAVETKDERPEQASNFYQQAEGHKLYKDLNPLS 69  
 Db 8 AKIASNVQKLTQAEKVLQKLGKADTKDEQEQCVQNFKNKLTGTRLQKDLRTYLA 67  
 QY 70 AVKVMHSHSKRVSSETLQEIYSSEWMDGHEELKAIWVNDLLWEDYEELADQAVRTWEIYV 129  
 Db 68 SVKAMHEASKKLNELQEVYEPWFGDEANKIAENNDLLWMDYHQKLVQALLTMDTYL 127  
 QY 130 AQSFEIKERIAKRGKLVYDSARHHLAVQNA-KDEAKTAKAEFEFNKQAVTFEDLNQ 188

Db 128 GQFPDIKSRKAGKLVYDSARHHLAVQNA-KDEAKTAKAEFEFNKQAVTFEDLNQ 187  
 QY 189 ELLEELPILYNSRICGYVTIFQNISNLRDVYR 221  
 Db 188 DLQBELPWSNRVGVYNTFQSIAGLEENFK 220  
 RESULT 3  
 US-08-652-972A-4  
 ; Sequence 4, Application US/08652972A  
 ; Patent No. 5723581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prndergast, George C.  
 ; APPLICANT: Sakamuro, Daitoku  
 ; TITLE OF INVENTION: Murine and Human Box-Dependent  
 ; TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P O Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/652,972A  
 ; FILING DATE: 24-MAY-1996  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/435,454  
 ; FILING DATE: 05-MAY-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bak, Mary E.  
 ; REGISTRATION NUMBER: 31,215  
 ; REFERENCE/DOCKET NUMBER: WST60BUSA  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-652-972A-4

Query Match 59.7%; Score 671.5; DB 1; Length 451;  
 Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
 Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;  
 QY 10 AGLEAKQVQKFSRAQEKVLQKLGAVETKDERPEQASNFYQQAEGHKLYKDLNPLS 69  
 Db 8 AKIASNVQKLTQAEKVLQKLGKADTKDEQEQCVQNFKNKLTGTRLQKDLRTYLA 67  
 QY 70 AVKVMHSHSKRVSSETLQEIYSSEWMDGHEELKAIWVNDLLWEDYEELADQAVRTWEIYV 129  
 Db 68 SVKAMHEASKKLNELQEVYEPWFGDEANKIAENNDLLWMDYHQKLVQALLTMDTYL 127  
 QY 130 AQSFEIKERIAKRGKLVYDSARHHLAVQNA-KDEAKTAKAEFEFNKQAVTFEDLNQ 188  
 Db 128 GQFPDIKSRKAGKLVYDSARHHLAVQNA-KDEAKTAKAEFEFNKQAVTFEDLNQ 187  
 QY 189 ELLEELPILYNSRICGYVTIFQNISNLRDVYR 221  
 Db 188 DLQBELPWSNRVGVYNTFQSIAGLEENFK 220

## RESULT 4

US-08-919-145-6  
 ; Sequence 6, Application US/08919145  
 ; Patent No. 5958753  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prendergast, George C.  
 ; TITLE OF INVENTION: Bau, A Binl Interacting Protein, and  
 ; TITLE OF INVENTION: Uses Therefor  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P.O. Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/919,145  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/025,482  
 ; FILING DATE: 29-AUG-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kodroff, Cathy A.  
 ; REGISTRATION NUMBER: 33,980  
 ; REFERENCE/DOCKET NUMBER: WST73AUSA  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-919-145-6

Query Match 59.7%; Score 671.5; DB 2; Length 451;  
 Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
 Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;  
 QY 10 AGLPAKOVKKFSRAQEKVLQKLGAVETKDERFEQASNFYQQOAGHKLKDLNPLS 69  
 Db 8 AGKIASNVQKLTTRAQEKVLQKLGKADETKDEQEQVQNFNQLTETGRLOKDLRTYLA 67  
 QY 70 AVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEKLAQAVRTWEIYV 129  
 Db 68 SVKAMHEASKKLNELQEVYEPDPGDEANKIAENNDLLWMDYHOKLVDQALLTMDTYL 127  
 QY 130 AQPSEIKERIAKGRKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKAQTVFEDLNQ 188  
 Db 128 GQFPDIKSRIAKGRKLVYDSARHYESLQAKKDEAKIAKAEELIKAKQVFEENVV 187  
 QY 189 ELLEELPLIYNSRIGCVVTIFQNISNLRDVFYR 221  
 Db 188 DLQELPLSLMNSRVGFYVNTFQSIAGLEENPHK 220

RESULT 5  
 US-08-919-145-6  
 ; Sequence 4, Application US/08870126  
 ; Patent No. 6048702  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prendergast, George C.  
 ; APPLICANT: Sakamuro, Daitoku  
 ; TITLE OF INVENTION: Murine and Human Box-Dependent  
 ; TITLE OF INVENTION: MYC-Interacting Protein (Binl) and Uses Therefor  
 ; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P O Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19477  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/870,126  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/435,454  
 ; FILING DATE: 05-MAY-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/652,972  
 ; FILING DATE: 24-MAY-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kodroff, Cathy A.  
 ; REGISTRATION NUMBER: 33,980  
 ; REFERENCE/DOCKET NUMBER: WST60CUSA  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-540-9200  
 ; TELEFAX: 215-540-5818  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 451 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-870-126-4

Query Match 59.7%; Score 671.5; DB 3; Length 451;  
 Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
 Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;  
 QY 10 AGLPAKOVKKFSRAQEKVLQKLGAVETKDERFEQASNFYQQOAGHKLKDLNPLS 69  
 Db 8 AGKIASNVQKLTTRAQEKVLQKLGKADETKDEQEQVQNFNQLTETGRLOKDLRTYLA 67  
 QY 70 AVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNDLLWEDYEKLAQAVRTWEIYV 129  
 Db 68 SVKAMHEASKKLNELQEVYEPDPGDEANKIAENNDLLWMDYHOKLVDQALLTMDTYL 127  
 QY 130 AQPSEIKERIAKGRKLVYDSARHLEAVQNA-KKDEAKTAKAEFEFNKAQTVFEDLNQ 188  
 Db 128 GQFPDIKSRIAKGRKLVYDSARHYESLQAKKDEAKIAKAEELIKAKQVFEENVV 187  
 QY 189 ELLEELPLIYNSRIGCVVTIFQNISNLRDVFYR 221  
 Db 188 DLQELPLSLMNSRVGFYVNTFQSIAGLEENPHK 220

RESULT 6  
 US-09-344-889-6  
 ; Sequence 6, Application US/09344889  
 ; Patent No. 6140465  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Prendergast, George C.  
 ; TITLE OF INVENTION: Bau, A Binl Interacting Protein, and  
 ; TITLE OF INVENTION: Uses Therefor  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Howson and Howson  
 ; STREET: Spring House Corporate Cntr, P.O. Box 457  
 ; CITY: Spring House  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA

ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,889  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/919,145  
FILING DATE:  
APPLICATION NUMBER: US 60/025,482  
FILING DATE: 29-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kodroff, Cathy A.  
REGISTRATION NUMBER: 33,980  
REFERENCE/DOCKET NUMBER: WST73AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-344-889-6

Query Match 59.7%; Score 671.5; DB 3; Length 451;  
Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;

QY 10 AGLFAKQVOKKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQAEGHKLKYNFLS 69  
DB 8 AGKIASNVOKLTRAQEKVLQKLGKADETKDEQFEQCVQNFKNQLTGTGLQKDLRYLA 67  
QY 70 ANKMHESKRVSETLOEIIYSSEWGHHEELKAIWNNDLLWEDYEEKLAQAVRTMEIYV 129  
DB 68 SVKAMEASKGLNECLQEVYEPDWPGRDEANKIAENNLLWMDYHQLVDQALLTMDTYL 127  
QY 130 AQPSEIKERIAKRGKLVVDYSARHLEAVONA-KKDEAKTAKAEFEFNKAQTVFEDLNQ 188  
DB 126 GQFPDIKSIARAKRGKLVVDYSARHHYESLQTKAKDEAKIAKAEELIKAKQVFEEMNV 187  
QY 189 ELLELPILYNSRIGCVYTFIFONISNLRDVYR 221  
DB 188 DLQELPSLWNSRVGYVNTFQSIAGLEENFHK 220

RESULT 7  
US-09-445-247-4  
Sequence 4, Application US/09445247  
Patent No. 6410238  
GENERAL INFORMATION:  
APPLICANT: Wistar Institute of Anatomy & Biology  
Prendergast, George C.  
Sakamuro, Daitoku  
TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/445,247  
FILING DATE: 03-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/870,126  
FILING DATE: 06-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60DPCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-445-247-4

Query Match 59.7%; Score 671.5; DB 4; Length 451;  
Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;

QY 10 AGLFAKQVOKKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQAEGHKLKYNFLS 69  
DB 8 AGKIASNVOKLTRAQEKVLQKLGKADETKDEQFEQCVQNFKNQLTGTGLQKDLRYLA 67  
QY 70 ANKMHESKRVSETLOEIIYSSEWGHHEELKAIWNNDLLWEDYEEKLAQAVRTMEIYV 129  
DB 68 SVKAMEASKGLNECLQEVYEPDWPGRDEANKIAENNLLWMDYHQLVDQALLTMDTYL 127  
QY 130 AQPSEIKERIAKRGKLVVDYSARHLEAVONA-KKDEAKTAKAEFEFNKAQTVFEDLNQ 188  
DB 126 GQFPDIKSIARAKRGKLVVDYSARHHYESLQTKAKDEAKIAKAEELIKAKQVFEEMNV 187  
QY 189 ELLELPILYNSRIGCVYTFIFONISNLRDVYR 221  
DB 188 DLQELPSLWNSRVGYVNTFQSIAGLEENFHK 220

RESULT 8  
PCT-US96-06231A-4  
Sequence 4, Application PC/JUS9606231A  
GENERAL INFORMATION:  
APPLICANT: Wistar Institute of Anatomy & Biology  
TITLE OF INVENTION: Murine and Human Box-Dependent  
TITLE OF INVENTION: Myc-Interacting Protein (Bin1) and Uses Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06231A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,454  
FILING DATE: 05-MAY-1995  
ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60APCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-06231A-4

Query Match 59.7%; Score 671.5; DB 5; Length 451;  
Best Local Similarity 61.5%; Pred. No. 5.4e-53;  
Matches 131; Conservative 32; Mismatches 49; Indels 1; Gaps 1;

QY 10 AGLFAKQVQKFSRAQKVLQKLGKAVETKDERFEQASNFYQQQAGHKLKDLKNFLS 69  
DB 8 AGKIASNVQKLTQAEKVLQKLGKADETKDEQEQCVQNFNKQLTEGTRLOKDLRTYLA 67

QY 70 AVKWHSSKRVSTLQEIYSSWDGHEELKAIWNNLLWEDYEKADQAVRTMEIY 129  
DB 68 SVKAMHEASKKNECLQEVYEPDPWGRDEANKIAENNDLLWMDYHQLVDQALLTMDTYL 127

QY 130 AQPSEIKERIAKGRKLVVDYDSARHLEAVONA-KKDEAKTAKAEFEFNKAQTVFEDLNQ 188  
DB 128 GQFPIKSRKAKGRKLVVDYDSARHLYESQTAKKDEAKIAKAEELIKAKRVFEEMNV 187

QY 189 ELLEELFILNSRIGCVYTFIFONISNLRDVFYR 221  
DB 188 DLOBELPSLNSRVGFYVNTFQSIAGLEENFHK 220

RESULT 9  
US-08-630-915A-24  
Sequence 24, Application US/08630915A  
Patent No. 6309820  
GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: HOFFMAN, No. 6309820h  
APPLICANT: KAY, Brian K.  
APPLICANT: FOWLER, Dana M.  
APPLICANT: MCCONNELL, Stephen J.  
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
NUMBER OF SEQUENCES: 227  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,915A  
FILING DATE: 03-APR-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 404 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-630-915A-24

Query Match 49.7%; Score 559.5; DB 4; Length 404;  
Best Local Similarity 59.1%; Pred. No. 7e-43;  
Matches 107; Conservative 30; Mismatches 43; Indels 1; Gaps 1;

QY 42 RFQSSASNFYQQQAGHKLKDLKNFLSAVKWHSSKRVSTLQEIYSSWDGHEELKA 101  
DB 8 RFQCVQNFNKQLTEGTRLOKDLRTYLASVKAMHEASKKNECLQEVYEPDPWGRDEANK 67

QY 102 IVNNLLWEDYEKADQAVRTMEIYVAQFSEIKERIAKGRKLVVDYDSARHLEAVON 161  
DB 68 IAEENDDLLWMDYHQLVDQALLTMDTYLGQFPDIKSRKAKGRKLVVDYDSARHLYESQT 127

QY 162 A-KKDEAKTAKAEFEFNKAQTVFEDLNQELLEELFILNSRIGCVYTFIFONISNLRDVFY 220  
DB 128 AKKDEAKTAKAEFEELIKAKRVFEEMNVLDQBELPSLNSRVGFYVNTFQSIAGLEENFH 187

QY 221 R 221  
DB 188 K 188

RESULT 10  
US-08-875-435B-3  
Sequence 3, Application US/08875435B  
Patent No. 6593304  
GENERAL INFORMATION:  
APPLICANT: Hasegawa, Kazuhide  
APPLICANT: Arakawa, Emi  
APPLICANT: Oda, Shoji  
APPLICANT: Matsuda, Yuzuru  
APPLICANT: Takahashi, Katsuhito  
APPLICANT: Sugahara, Michihiro  
APPLICANT: Ishiyama, Haruo  
TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING  
TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO  
TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND  
TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE  
TITLE OF INVENTION: RECOMBINANT DNA  
FILE REFERENCE: 07898-013001  
CURRENT APPLICATION NUMBER: US/08/875,435B  
CURRENT FILING DATE: 1997-07-25  
PRIOR APPLICATION NUMBER: PCT/JP96/00134  
PRIOR FILING DATE: 1996-01-25  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 1972  
TYPE: PRT  
ORGANISM: Mus musculus  
US-08-875-435B-3

Query Match 10.0%; Score 113; DB 4; Length 1972;  
Best Local Similarity 24.8%; Pred. No. 0.23;  
Matches 57; Conservative 38; Mismatches 97; Indels 38; Gaps 9;

QY 2 AEGKAGGAA---GLFAKQVQKFSRAQKVLQKLGKAVETKDERFEQASNFY----- 51  
DB 1299 AEGKAIKLAQVASJGSQLODTQELLQETROKLN--VSTKLQLEDERNSLQLODEEM 1356

QY 52 --QQQAGH-----KLYKDLKNFLSAVKWHSSKRVSTLQEIYSSWDGHEELKA 97  
DB 1357 EAKQNLERHVSITNLQSDSKKQLQDFASTIEVMEGKRLQKEMEGLSQQYBEKAAAYD 1416

QY 98 ELKAIWNNLLWEDYEKADQAVRTMEIYVAQFSEIKERIAKGRKLVVDYDS 152

Db 1417 KLEK---TKNRLQQLDLDVLDNQRLVSNLEKKKKFQDL---LAEKNTSSKYADE 1470  
 QY 153 RHLEAVQNAKDEA-KTAKAEFEFNKAQTVFEDLNQELLEELPILYNSR 201  
 Db 1471 RDAEAREKETKALSARALEEALAEKEELERTNKMLKAEMEDLVSSK 1520

## RESULT 11

US-09-150-867-1  
 ; Sequence 1, Application US/09150867  
 ; Patent No. 6645748  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wood, Kenneth W.  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Goldstein, Lawrence S.B.  
 ; APPLICANT: Cleveland, Don W.  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for  
 ; FILE REFERENCE: 18557C-000110US  
 ; CURRENT APPLICATION NUMBER: US/09/150.867  
 ; EARLIER FILING DATE: 1998-09-10  
 ; EARLIER APPLICATION NUMBER: US 60/058,645  
 ; EARLIER FILING DATE: 1997-09-11  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2954  
 ; TYPE: PRT  
 ; ORGANISM: Xenopus sp.  
 ; FEATURE:  
 ; OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)  
 ; OTHER INFORMATION: member of the kinesin superfamily of microtubule  
 ; OTHER INFORMATION: motor proteins  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (1)..(472)  
 ; OTHER INFORMATION: kinesin like motor domain  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (473)..(2752)  
 ; OTHER INFORMATION: rod domain  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (2753)..(2954)  
 ; OTHER INFORMATION: tail domain  
 ; US-09-150-867-1

Query Match 9.9%; Score 111.5; DB 4; Length 2954;  
 Best Local Similarity 22.1%; Pred. No. 0.54;  
 Matches 57; Conservative 46; Mismatches 88; Indels 67; Gaps 10;  
 QY 13 FAKVQKFSRAQEKVLQKLGKAVETKDERPEQASNFYQQAEGHKLAVK 72  
 Db 1720 FNEAQQKTTKEQ-----CLINENKELEQSC---HRLQCEIEBELMKSLKDSALE 1767  
 QY 73 VMHESKRVSETLQEIYSSWDGHEELK---ATVWNNDLWEDY----- 113  
 Db 1768 TLKESQKVINLQEMEMWLE-MEELKNSORTVIAERDQLQDLRESVEMSTETQDDL 1826  
 QY 114 --EKLQADQVRTWEIYVAQFSEIKERIAKGRKLV-----DYDSARHIL- 156  
 Db 1827 KAQEAQQQKQVQEL-TSQISVLQEKISLLENQNVATVKTSLRDLNQSKQLHF 1885  
 QY 157 -----EAVQNAKDEAKTA-KAEFEFNKAQTVFEDLNQELLEELPILY--N 199  
 Db 1886 SEIETLSLSKEKEFALQAEKQKADAAKTIDITEXISNIEBQLQQAATNLKETLYERE 1945  
 QY 200 SRIGCVTVIFQNISLRD 217  
 Db 1946 SLIQCKEQLAINTHLRE 1963

## RESULT 12

US-08-533-306A-6  
 ; Sequence 6, Application US/08533306A  
 ; Patent No. 5837457  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Pu  
 ; APPLICANT: Collins, Francis S.  
 ; APPLICANT: Siciliano, Michael J.  
 ; APPLICANT: Claxton, David  
 ; TITLE OF INVENTION: Markers for Detection of Chromosome 16  
 ; TITLE OF INVENTION: Rearrangements  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Harnes, Dickey & Pierce, P.L.C.  
 ; STREET: P.O. Box 828  
 ; CITY: Bloomfield Hills  
 ; STATE: MI  
 ; COUNTRY: USA  
 ; ZIP: 48303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/533.306A  
 ; FILING DATE: September 25, 1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Smith, Deann F.  
 ; REGISTRATION NUMBER: 36693  
 ; REFERENCE/DOCKET NUMBER: 2115-00869COB  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (810) 641-1600  
 ; TELEFAX: (810) 641-0270  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 816 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-533-306A-6

Query Match 9.8%; Score 110; DB 2; Length 816;  
 Best Local Similarity 22.9%; Pred. No. 0.13;  
 Matches 52; Conservative 39; Mismatches 104; Indels 32; Gaps 6;  
 QY 2 AEGKAGGAA---GLFAKQVKFSPRAQEKVLQKLGKAVETKDERPEQASNFYQQAEGH 58  
 Db 177 AEGKAIKADVASLASQLQDTQELLQETRQKLN--VSTKLRQLEERNLSLQDLDEEM 234  
 QY 59 KLYKDLKNFLSAVKVMHESKRVSETLQEIYSSWDGHEELKAIYVNNNDLLWEDYEKLA 118  
 Db 235 EAKQNLHERHISTNLIQLSDSKKQLQDPASTVEALBEGKKRFQKEIEN---LTQVEEKAA 291  
 QY 119 -----DQAVRTWEIYVAQFSEIKERIAKGRKL-----VDYDSARH 155  
 Db 292 AYDKLETKNRLQQLDLDVLDNQRLVSNLEKKKKFQDLQDLAEKKNISSKYADERDR 351  
 QY 156 LEAVQNAKDEA-KTAKAEFEFNKAQTVFEDLNQELLEELPILYNSR 201  
 Db 352 AEAREKETKALSARALEEALAEKEELERTNKMLKAEMEDLVSSK 398

## RESULT 13

US-08-742-923A-6  
 ; Sequence 6, Application US/08742923A  
 ; Patent No. 5869611  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Pu  
 ; APPLICANT: Collins, Francis S.  
 ; APPLICANT: Siciliano, Michael J.



```
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
; FILING DATE: No. 5869611ember 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-923A-6

Query Match          9.8%; Score 110; DB 2; Length 816;
Best Local Similarity 22.9%; Pred. No. 0.13;
Matches 52; Conservative 39; Mismatches 104; Indels 32; Gaps 6;

QY 2 AEGKAGGAA---GLFAKQVQKFSRAQEKVLQKLGKAVETKDRFQSGASNFYQQQAEGH 58
|||||
Db 177 AEGKAIKLAKDVASLSQQTQELLQETQKLN--VSTKLQLEERNSLQDQDEEM 234
|||||
QY 59 KLYDKLKNFLSAVKVHSHSKRVSETLQEIYSSEWDGHEELKAIWNNDDLWEDYEKLA 118
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 EAKQNLERHISTNLQLSDSKKLQDPASTVLEAEEGKGRFQKEIEN---LTQQYEEKAA 291
|||||
QY 119 -----DQAVRTWEIYVAQFSEIKERIAKGRKL-----VDYDSARHH 155
|||||
Db 292 AYDKLETKNRLQELDDLVDLDNQRQLVSNLEKQKQKQKFDQLLAEKNISSKYADERDR 351
|||||
QY 156 LEAVQNAKKDEA-KTAKAEFEFNKAQTVFEDLNQELLEELPILYNR 201
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 AEAEAREKETKALSARALEALEAKEELERTNKMLKAEMEDLVSK 398
|||||

RESULT 14
US-08-533-306A-4
; Sequence 4, Application US/08533306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
```

```
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533,306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869COB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-533-306A-4

Query Match          9.8%; Score 110; DB 2; Length 885;
Best Local Similarity 22.9%; Pred. No. 0.14;
Matches 52; Conservative 39; Mismatches 104; Indels 32; Gaps 6;

QY 2 AEGKAGGAA---GLFAKQVQKFSRAQEKVLQKLGKAVETKDRFQSGASNFYQQQAEGH 58
|||||
Db 246 AEGKAIKLAKDVASLSQQTQELLQETQKLN--VSTKLQLEERNSLQDQDEEM 303
|||||
QY 59 KLYDKLKNFLSAVKVHSHSKRVSETLQEIYSSEWDGHEELKAIWNNDDLWEDYEKLA 118
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 304 EAKQNLERHISTNLQLSDSKKLQDPASTVLEAEEGKGRFQKEIEN---LTQQYEEKAA 360
|||||
QY 119 -----DQAVRTWEIYVAQFSEIKERIAKGRKL-----VDYDSARHH 155
|||||
Db 361 AYDKLETKNRLQELDDLVDLDNQRQLVSNLEKQKQKQKFDQLLAEKNISSKYADERDR 420
|||||
QY 156 LEAVQNAKKDEA-KTAKAEFEFNKAQTVFEDLNQELLEELPILYNR 201
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 AEAEAREKETKALSARALEALEAKEELERTNKMLKAEMEDLVSK 467
|||||

RESULT 15
US-08-742-923A-4
; Sequence 4, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
```

```
; FILING DATE: No. 5869611ember 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-923A-4

Query Match          9.8%; Score 110; DB 2; Length 885;
Best Local Similarity 22.9%; Pred. No. 0.14;
Matches 52; Conservative 39; Mismatches 104; Indels 32; Gaps 6;

QY 2 AEGKAGGAA---GLFAKQVQKFSRAQEKVLQKLGKAVETDERFPOSASNFYQQAEHG 58
Db 246 AEGKAIKLAKDVASLSQQLQELQEBETRQKLN--VSTKLRQLEERNLSQDLDEEM 303
QY 59 KLYDILKNFLSAVKVWHSSKRVSETLQEIYSSSEWDGHEELKATVWNNDLLWEDYEKLA 118
Db 304 EAKQNLERHISTINIQLSDSKKLQDPASTVTEALEEGKKRFQKEIEN---LTQOYEKAA 360
QY 119 -----DQAVRTWEIYVAQFSEIKERIAKGRKL-----VDYDSARHH 155
Db 361 AYDKLEKTKNRLQQLDLDIVDLNQROLVSNLEKKQKFDQLAEEKNISSKYADERDR 420
QY 156 LEAVQNAKKDEA-KTAKAEERFNAQTVPEDLNQELLEELPILYNSR 201
Db 421 ABABAREKETKALSRLALEALEAKELELRTNKMKAEMEDLVSSK 467
```

Search completed: March 4, 2004, 17:32:27  
Job time : 10.2904 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:23:55 ; Search time 10.135 Seconds  
(without alignments)  
2097.510 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_1\_221

Perfect score: 1125  
Sequence: 1 MAEGKAGAGLPAKQVQKK.....IGCVYTFQINSLRDVFR 221

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	683	60.7	482	2 JC5593	amphiphysin Ii2 -
2	620.5	55.2	695	2 S62400	amphiphysin (clone
3	618.5	55.0	682	1 S22700	amphiphysin - chic
4	405	36.0	461	2 T22946	hypothetical prote
5	194.5	17.3	270	2 T43000	RVS161 protein hom
6	192.5	17.1	264	2 T40661	Yeast reduced viab
7	179.5	16.0	265	1 S19516	RVS161 protein - y
8	158.5	14.1	466	2 T11684	RVS167 protein hom
9	154.5	13.7	482	1 S40887	RVS167 protein - y
10	129.5	11.5	1992	2 A47297	myosin heavy chain
11	127	11.3	377	2 S52537	eml L 15 protein -
12	120.5	10.7	2116	2 A26655	myosin heavy chain
13	118	10.5	399	2 E71169	hypothetical prote
14	116	10.3	1940	2 A59287	myosin heavy chain
15	116	10.3	2401	2 T28676	rhostry protein -
16	116	10.3	3225	2 I52300	giantin - human
17	116	10.3	3259	1 A56539	giantin - human
18	115	10.2	527	2 S33068	myosin heavy chain
19	114.5	10.2	532	2 S54871	M protein - Strept
20	114	10.1	226	2 F71108	hypothetical prote
21	114	10.1	1061	2 D98008	conserved hypothet
22	114	10.1	1078	2 T18352	protein P120 - Myc
23	113.5	10.1	1087	2 T30330	Gelsolin-related p
24	113	10.0	1005	2 A64465	hypothetical prote
25	113	10.0	1300	2 T18364	ro-3 protein - Neu
26	113	10.0	1938	2 JC5421	smooth muscle myos
27	113	10.0	1972	2 JC5420	smooth muscle myos
28	112.5	10.0	1138	2 T24635	hypothetical prote
29	112.5	10.0	2245	2 T18278	myosin heavy chain

30	112	10.0	600	2 A45112	major parafagella
31	112	10.0	2007	1 B43402	myosin heavy chain
32	111.5	9.9	1790	2 S67593	transport protein
33	111.5	9.9	2954	2 T14156	kinesin-related pr
34	110.5	9.8	1169	2 A64505	P115 homolog - Met
35	110.5	9.8	1491	2 D82166	cell division prot
36	110	9.8	407	2 S23325	M2 protein precurs
37	110	9.8	841	2 A86188	hypothetical prote
38	110	9.8	1640	2 D86798	prophage p13 prote
39	109.5	9.7	944	2 S26710	spindle pole body
40	109.5	9.7	1225	2 A49464	chromosome segrega
41	109	9.7	281	2 F75216	hypothetical prote
42	109	9.7	520	2 F70350	recombination prot
43	109	9.7	736	2 D86271	protein F16A14.2
44	109	9.7	942	2 A64379	isoleucine-tRNA li
45	109	9.7	1170	2 A56157	chromosome segrega

## ALIGNMENTS

### RESULT 1

JC5593  
amphiphysin Ii2 - human  
C:Species: Homo sapiens (man)  
C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 09-Jun-2000  
C:Accession: JC5593  
R:Tsutsui K.; Maeda, Y.; Tsutsui, K.; Seki, S.; Tokunaga, A.  
Biochem. Biophys. Res. Commun. 236, 178-183, 1997  
A:Title: cDNA cloning of a novel amphiphysin isoform and tissue-specific expression of i  
A:Reference number: JC5593; MUID:97366618; PMID:9223448  
A:Accession: JC5593  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-482 <TSU>  
A:Cross-references: DDBJ:AF001383; MID:g2199534; PIDN:AAB61363.1; PID:g2199535  
A:Experimental source: fetal brain  
C:Comment: This protein is involved in the synaptic vesicle recycling and in the regulat  
C:Superfamily: amphiphysin; RVS161 protein homology  
F:16-275/Domain: RVS161 protein homology <RVS>  
F:410-481/Domain: SH3 #status predicted <SH3>

Query Match 60.7%; Score 683; DB 2; Length 482;

Best Local Similarity 61.0%; Pred. No. 4.2e-35;

Matches 136; Conservative 32; Mismatches 53; Indels 2; Gaps 2;

Qy	1	MAE-GRAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDRFQSSASNFYQQQAGHK	59
Db	1	MAEGRSGKVTAGKIANVQKGLTRAQEKVLQKLGKADETKDRFQSQVFNKQLTEGTR	60
Qy	60	LYKDLNLFSAVKVMHSSKRVSETLQEIYSEWDGHEELKAIWNNDLLWEDYEELAD	119
Db	61	LQKDLRTYLSVFAVMEASKNLCEIQEYVEPDPGRDEANKIAENNDLLWMDYHQKLV	120
Qy	120	QAVRTWEIVQAFSEIKERIAKGRKLVQYDSARHLEAVQNA-KKDRAKTAKAEENK	178
Db	121	QALTMDTYLGQFPDKISRIKGRKLVQYDSARHYSLSQPAKKDKAKIAEELIK	180
Qy	179	AQTVFEDNLQELLEELPILYNSRIGCYVTIFQINSLRDVFR	221
Db	181	AQKVFEMNVDLQELPLWSNRSGVYVYVTFQSIAGLENFHK	223

### RESULT 2

S62400  
amphiphysin (clone 22-2) - human  
C:Species: Homo sapiens (man)  
C:Date: 12-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999  
C:Accession: S62400; I37166  
R:David, C.; Solimena, M.; de Camilli, P.  
FEBS Lett. 351, 73-79, 1994  
A:Title: Autoimmunity in Stiff-Man Syndrome with breast cancer is targeted to the C-term  
A:Reference number: S48686; MUID:94357284; PMID:8076697

A:Accession: S62400  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-695 <DAV>  
A:Cross-references: EMBL:U07616; NID:G550449; PIDN:AAA21865.1; PID:G550450  
R:Yamamoto, R.; Li, X.; Winter, S.; Francke, U.; Kilmann, M.W.  
Hum. Mol. Genet. 4, 265-268, 1995  
A:Title: Primary structure of human amphiphysin, the dominant autoantigen of paraneoplastic cerebellar degeneration  
A:Reference number: 137166; MUID:95276740; PMID:7757077  
A:Accession: 137166  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-695 <RES>  
A:Cross-references: EMBL:X81438; NID:G662991; PIDN:CAA57197.1; PID:G662992  
C:Genetics:  
A:Gene: GDB:AMPH  
A:Cross-references: GDB:386990  
A:Map position: 7p4-7p13  
C:Superfamily: amphiphysin; RVS161 protein homology  
F:11-270/Domain: RVS161 protein homology <RVS>  
Query Match 55.2%; Score 620.5; DB 2; Length 695;  
Best Local Similarity 54.7%; Pred. No. 4.6e-31;  
Matches 116; Conservative 49; Mismatches 46; Indels 1; Gaps 1;  
QY 11 GLFAKQVQKFRQAEKVLQKLGKAVETKDERFEQSASNFYQQQAGHKLKYLKLNFLSA 70  
DB 7 GIFAQNVQKRLNRAQEKVLQKLGKADETKDEQFEYVQNFQKQAEQTRQLQRELGLAA 66  
QY 71 VKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNLLWEDYEKLADQAVTMEIYVA 130  
DB 67 IKGMEASQKLSTESLHEVYEPDYGREDVQKMGKCVLMDFHQKLVGSLTLDLYLG 126  
QY 131 QPSEIKERIAKGRKLVYDVSARHLEAVQNAK-KDEAKTAAEEEFNKQATVFEDLNQE 189  
DB 127 QPFDIKNRIAKESKRLVYDVSARHLEALQSSKRSRISKAEERFQKQKVFEEFNVD 186  
QY 190 LLEELPILNSRIGCYVTFIFONISNLRDVF 221  
DB 187 LOELPFLWSRVGVYVNTFNKVSLEAKFHK 218  
RESULT 3  
S22700  
amphiphysin - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S22700  
R:Lichte, B.; Veh, R.W.; Meyer, H.E.; Kilmann, M.W.  
EMBO J. 11, 2521-2530, 1992  
A:Title: Amphiphysin, a novel protein associated with synaptic vesicles.  
A:Reference number: S22700; MUID:92331604; PMID:1628617  
A:Accession: S22700  
A:Molecule type: mRNA  
A:Residues: 1-682 <LIC>  
A:Cross-references: EMBL:X60422; NID:G62842; PIDN:CAA42953.1; PID:G62843  
C:Superfamily: amphiphysin; RVS161 protein homology  
F:11-270/Domain: RVS161 protein homology <RVS>  
Query Match 55.0%; Score 618.5; DB 1; Length 682;  
Best Local Similarity 54.1%; Pred. No. 5.9e-31;  
Matches 120; Conservative 47; Mismatches 50; Indels 5; Gaps 2;  
QY 1 MAEGKAGGAGAFQKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQQAGHKL 60  
DB 1 MADMK-----TGIFAKNVQKRLNRAQEKVLQKLGKADETKDEQFEYVQNFQKQAEGRSL 56  
QY 61 YKDLKLNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNLLWEDYEKLADQ 120  
DB 57 QRELRAVLAIAKGMQDASKLTESLHEVYEPDYGREDVQKMGKCVLMDFHQKLVG 116  
QY 121 AVRTMEIYVAQSEIKERIAKGRKLVYDVSARHLEAVQNAK-KDEAKTAAEEEFNK 179

Db 117 SLLTLDTYLQFPDIKTRIAKRSKRLVDYDSARHLEALQSSKRSRISKAEERFQKA 176  
QY 180 QTVFEDLNQELLELPILNSRIGCYVTFIFONISNLRDVFYR 221  
Db 177 QKVFEFETDQLQELPSLWSRVGVYVNTFNKVSLEAKFHK 218  
RESULT 4  
T22946  
hypothetical protein F58G6.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000  
C:Accession: T22946  
R:Lloyd, C.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: Z19641  
A:Accession: T22946  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-461 <WIL>  
A:Cross-references: EMBL:Z68217; PIDN:CAA92465.1; GSPDB:GN00022; CESP:F58G6.1  
C:Genetics:  
A:Gene: CESP:F58G6.1  
A:Map position: 4  
A:Introns: 19/3; 57/3; 96/3; 130/3; 157/3; 186/3; 219/3; 271/3; 302/3; 335/3; 357/3  
C:Superfamily: amphiphysin; RVS161 protein homology  
Query Match 36.0%; Score 405; DB 2; Length 461;  
Best Local Similarity 39.6%; Pred. No. 5.7e-18;  
Matches 84; Conservative 52; Mismatches 74; Indels 2; Gaps 2;  
QY 10 AGLFAKQVQKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQQAGHKLKYLKLNFLS 69  
DB 2 ADLFNKHLKATNRTKELLEGIGKAKATQDEVDFQHAANLNKQSKSEKLNKDVKNVSS 61  
QY 70 AVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNLLWEDYEKLADQAVTMEIYV 129  
DB 62 ALRTLLSAEQRLDTRDAYEPEDREHLTAIFDNLDIQTNELEKTVCDLFPQTVTQV 121  
QY 130 AQFSEIKERIAKGRKLVYDVSARHLEAVQ-NAKID-EAKTAAEEEFNKQATVFEDLN 187  
DB 122 NQFPDLAKKIEKGRKLVYDVSASNSFNSVKASSKNDPKLAKATMEIQAEEQVMEYN 181  
QY 188 QELLELPILNSRIGCYVTFIFONISNLRDVF 219  
DB 182 NELLEILPAVDFSRITFFVDTLQTLFNANSVY 213  
RESULT 5  
T43000  
RVS161 protein homolog - fission yeast (Schizosaccharomyces pombe) (fragment)  
C:Species: Schizosaccharomyces pombe  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T43000  
R:Yoshioka, S.; Kato, K.; Nakai, H.; Okayama, H.; Nojima, H.  
DNA Res. 4, 363-369, 1997  
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.  
A:Reference number: Z17323; MUID:98162722; PMID:9501991  
A:Accession: T43000  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-270 <YOS>  
A:Cross-references: EMBL:D89200; NID:G1749607; PIDN:BAAL3861.1; PID:G1749608  
A:Experimental source: strain PR745  
C:Superfamily: RVS161 protein; RVS161 protein homology  
F:10-261/Domain: RVS161 protein homology <RVS>  
Query Match 17.3%; Score 194.5; DB 2; Length 270;  
Best Local Similarity 29.0%; Pred. No. 3e-05;  
Matches 56; Conservative 33; Mismatches 93; Indels 11; Gaps 5;  
QY 18 QKFSRAQEKVLQKLGKAVETKDERFEQSASNFYQQQAGHKLKYLKLNFLSAVKVMHES 77

```

Db      13 KAVNRAGTSVMKTHGCHVETVDFEFTEERRRYHTMESAAKQKQKAGVLDALRAMTAS 72
QY      78 SKRVSETLOEYIY--SSEWDGHEELKAIWVNDLLWEDYEKLQADQAVRTMEI-----YVA 130
Db      73 QTRIANITDAFYGDAGSGDGVSAAYRLV-VKDLADATVKE--LDGPFRTTVDLPISRFC 129
QY      131 OFSEIKERIAKRGKLVYDSARHLE-AVQNAKKDEAKTAKAESEFNKAQTVPFDLNOE 189
Db      130 YFPDINAATIKRNHKLHDHADRKAQVQLVDKPSNDTTKLPRTKEGAAMAKEVYETLNNQ 189
QY      190 LLEELPILYNSRI 202
Db      190 LVSELPOLIALRV 202

RESULT 6
T40661
Yeast reduced viability upon starvation protein 161 homolog, implicated in cell growth
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C:Accession: T40661
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
Submitted to the EMBL Data Library, November 1998
A:Reference number: Z21889
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: T40661
A:Molecule type: DNA
A:Residues: 1-264 <LYN>
A:Cross-references: EMBL:AL034352; PIDN:CAA22181.1; GSPDB:GN00067; SPDB:SPBC725.09c
A:Experimental source: strain 972h-; cosmid c725
C:Genetics:
A:Gene: SPDB:SPBC725.09c
A:Map position: 2
A:Introns: 16/3
C:Superfamily: RVS161 protein; RVS161 protein homology
F:4-255/Domain: RVS161 protein homology <RVS>

Query Match      17.1%; Score 192.5; DB 2; Length 264;
Best Local Similarity 27.8%; Pred. No. 3.9e-05;
Matches 54; Conservative 35; Mismatches 92; Indels 13; Gaps 5;

QY      18 OKKPSRAQEKVLQKLGKAVETKDERPEQSASNFYQQAEGHLYKDLKNFLSAVKVMEES 77
Db      7 KAVNRAGTSVMKTHGCHVETVDFEFTEERRRYHTMESAAKQKQKAGVLDALRAMTAS 66
QY      78 SKRVSETLOEYIYSSSEWDGHEELKAIWVNDLLWEDYEE---KLADQAVRTMEI-----YV 129
Db      67 QTRIANITDAFYGDA--GSKDGVSAAYRQ--VVEDLDADTVKELDGFRTTVDLPISRFC 122
QY      130 AQPSEIKERIAKRGKLVYDSARHLE-AVQNAKKDEAKTAKAESEFNKAQTVPFDLNOE 188
Db      123 SYFPDINAATIKRNHKLHDHADRKAQVQLVDKPSNDTTKLPRTKEGAAMAKEVYETLNN 182
QY      189 ELLELPILYNSRI 202
Db      183 QLVSELPOLIALRV 196

RESULT 7
S19516
RVS161 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YCR009c
C:Species: Saccharomyces cerevisiae
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 21-Jul-2000
R:Goffeau, A.; Purnelle, B.; Skala, J.
Submitted to the Protein Sequence Database, March 1992
A:Reference number: S19420
A:Accession: S19516
A:Molecule type: DNA
A:Residues: 1-265 <GOF>
A:Cross-references: EMBL:X59720; NID:g1907116; PIDN:CAA42326.1; PID:g1907152; GSPDB:GN00

```

```

R:Crouzet, M.; Urdaci, M.; Dulau, L.; Aigle, M.
Yeast 7, 727-743, 1991
A:Title: Yeast mutant affected for viability upon nutrient starvation: characterization
A:Reference number: S28651; MUID:92133163; PMID:1776363
A:Accession: S28651
A:Molecule type: DNA
A:Residues: 1-94, 'E', 96-265 <CRO>
A:Cross-references: EMBL:X63315; NID:g4417; PIDN:CAA44926.1; PID:g4418
R:Urdaci, M.; Dulau, L.; Aigle, M.; Crouzet, M.
Yeast 6, 173-176, 1990
A:Title: Sequence of the yeast gene RVS161 located on chromosome III.
A:Reference number: S22849; MUID:90224366; PMID:2183524
A:Accession: S22849
A:Molecule type: DNA
A:Residues: 1-94, 'E', 96-265 <URD>
A:Cross-references: EMBL:X63315; NID:g4417; PIDN:CAA44926.1; PID:g4418
R:Skala, J.; Purnelle, B.; Goffeau, A.
Yeast 8, 409-417, 1992
A:Title: The complete sequence of a 10.8 kb segment distal of SUF2 on the right arm of c
K genes.
A:Reference number: S25353; MUID:92327849; PMID:1626432
A:Contents: annotation
C:Genetics:
A:Gene: SGP:RVS161; MIPS:YCR009c
A:Cross-references: SGD:S000602; MIPS:YCR009c
A:Map position: 3R
C:Superfamily: RVS161 protein; RVS161 protein homology
F:4-257/Domain: RVS161 protein homology <RVS>

Query Match      16.0%; Score 179.5; DB 1; Length 265;
Best Local Similarity 24.1%; Pred. No. 0.00025;
Matches 52; Conservative 50; Mismatches 87; Indels 27; Gaps 6;

QY      18 OKKPSRAQEKVLQKLGKAVETKDERPEQSASNFYQQAEGHLYKDLKNFLSAVKVMEES 77
Db      7 KAVNRAGTSVMKTHGCHVETVDFEFTEERRRYHTMESAAKQKQKAGVLDALRAMTAS 64
QY      78 SKRVSETLOEYIY--SSEW-----DGHEELKAIWVNDLLWEDYEEK-----LAQQA 121
Db      65 QTTIAEVISNLYDSDSKVAGGVNNGVYVLCV-----QDFDSETVKQLDGLPRTV 116
QY      122 VRTMEIYVAQFSEIKERIAKRGKLVYDSARHLL-EAVQNAKKDEAKTAKAESEFNKAQ 180
Db      117 LDPITFTSTFYKFEIEAIXGRDHKKQDFDAKAKVRLVDRKPAKASKLPRAEKLSLAK 176
QY      181 TVFDLNOELLELPILYNSRIGCYVTIFONTISNLR 216
Db      177 DIFENLNNQKTELPQLVSLRVPYDFPSPFALIKIQ 212

RESULT 8
T11684
RVS167 protein homolog - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: T11684
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, September 1998
A:Reference number: Z17313
A:Accession: T11684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-466 <SEE>
A:Cross-references: EMBL:AL031536
A:Experimental source: strain 972h(-)
C:Genetics:
A:Map position: IIR
A:Introns: 444/3
A:Note: SPBC21D10.12
C:Superfamily: RVS167 protein; RVS161 protein homology; SH3 homology
F:4-270/Domain: RVS161 protein homology <RVS>
F:414-463/Domain: SH3 homology <SH3>

```

Query Match 14.1%; Score 158.5; DB 2; Length 466;  
 Best Local Similarity 23.8%; Pred. No. 0.0034;  
 Matches 57; Conservative 39; Mismatches 99; Indels 45; Gaps 7;

QY 15 KQVKKFSRAQKVLQKAVETKDERFEQASNFYQQAEGHKLKYLKXFLSAVKVM 74  
 DB 4 KGFALKARPTQRLSKFNVNITKPIYEDAGRRFKSLTEAKKLAEDAKKYTDALNGL 63

QY 75 HESSKRVSEYLOIY-----SSWDGH-----BELKAIWNNDLLMEDYEK 116  
 DB 64 LNHQIGFADACIEIYXIPISGRASDPSEYQEGNAGIEAAKYEIV-----DIQKN 116

QY 117 LADQAVTMEIYVAAF-----SEIKERIAKRGKLVYDSARHHLAEVQNAK-- 163  
 DB 117 LASE-----MDVINTRVNTPELLKIVKDVDDKLLKRDHKQLDYDRHRSFKEQKDK 172

QY 164 --KDEAKTAKABEEFNKAQTVFEDLNQELLELPILYNSRGYVYIFQNTLSNLR-DVFY 220  
 DB 173 SLDEKCLYEATFAFQSSQSEYIYNNMLKEELPKLFALASFIAPLFGGYNNLVY 232

RESULT 9  
 S40887  
 RVS167 protein - Yeast (Saccharomyces cerevisiae)  
 N;Alternate names: protein YDR388w  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 31-Mar-1992 #sequence\_revision 06-Feb-1995 #text\_change 21-Jul-2000  
 C;Accession: S40887; S69672  
 R;Bauer, F.; Urdaci, M.; Aigle, M.; Crouzet, M.  
 Mol. Cell. Biol. 13, 5070-5084, 1993  
 A;Title: Alteration of a yeast SH3 protein leads to conditional viability with defects in  
 A;Reference number: S40887; MUID:93330299; PMID:836735  
 A;Accession: S40887  
 A;Molecule type: DNA  
 A;Residues: 1-482 <BAU>  
 A;Cross-references: EMBL:M92092; NID:g172615; PIDN:AAA35051.1; PID:g172616  
 R;Dietrich, F.S.  
 submitted to the EMBL Data Library, July 1995  
 A;Description: The sequence of S. cerevisiae cosmid 9481, 9509, 9926, 9461, and lambda  
 A;Reference number: S69665  
 A;Accession: S69672  
 A;Molecule type: DNA  
 A;Residues: 1-482 <DIE>  
 A;Cross-references: EMBL:U32274; NID:g927313; PIDN:AAB4830.1; PID:g927321; GSPDB:GN0000  
 C;Genetics:  
 A;Gene: SGD:RVS167; MIPS:YDR388w  
 A;Cross-references: SGD:S0002796; MIPS:YDR388w  
 A;Map position: 4R  
 C;Superfamily: RVS167 protein; RVS161 protein homology; SH3 homology  
 C;Keywords: transmembrane protein  
 F;4-270/Domain: RVS161 protein homology <RVS>  
 F;292-422/Region: alanine/glycine/proline-rich  
 F;428-477/Domain: SH3 homology <SH3>

Query Match 13.7%; Score 154.5; DB 1; Length 482;  
 Best Local Similarity 24.4%; Pred. No. 0.017;  
 Matches 61; Conservative 38; Mismatches 100; Indels 51; Gaps 8;

QY 15 KQVKKFSRAQKVLQKAVETKDERFEQASNFYQQAEGHKLKYLKXFLSAV--- 71  
 DB 4 KGFALKARAPGFRQKFMGQGTEDPVYEDARRPQLEQETKLSSEKRYSTAVNGM 63

QY 72 -----KVMHSSKRV-----ETLQEIYSSWDGHBEELKAIWVN-NDLLWED--- 112  
 DB 64 LTHQIGFAKSMEEIFPKSGKSDPNATIPEDNPQIEASQYRAIVAELOETLKPDLAL 123

QY 113 YEELKADQAVTMEIYVAAFSEIKERIAKRGKLVYDSARHHLAEVQNAKDEAKTAKA 172  
 DB 124 VEEKIYVTPQOELLKI-----ITYRKQATKRNHKKDLDL---RHLNTYNNKHEKKEPTAKD 176

QY 173 EEEFNKAQTVFEDLNQELLELPIL-----YNSRIGCVYTFIQ 210  
 DB 177 EERLYRAQAVEVAQGEYDYNDLKTQLPILFSLAEAFVPLFVSFFYFMQLNIFYLYN 236

QY 211 NISNLRDVY 220  
 DB 237 RLQDMKPIYF 246

RESULT 10  
 A47297  
 myosin heavy chain form B, nonmuscle - African clawed frog  
 C;Species: Xenopus laevis (African clawed frog)  
 C;Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 02-Feb-2001  
 C;Accession: A47297; A55441  
 R;Bhatia-Dey, N.; Adelstein, R.S.; David, I.B.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 2856-2859, 1993  
 A;Title: Cloning of the cDNA encoding a myosin heavy chain B isoform of Xenopus nonmuscle  
 A;Reference number: A47297; MUID:93219383; PMID:8464900  
 A;Accession: A47297  
 A;Status: preliminary; not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-1992 <BHA>  
 A;Cross-references: GB:L09740; NID:g214623; PIDN:AAA49915.1; PID:g214624  
 A;Experimental source: XTC cells  
 A;Note: sequence extracted from NCBI backbone (NCBI:P128722)  
 R;Kelley, C.A.; Oberman, F.; Yisraeli, J.K.; Adelstein, R.S.  
 J. Biol. Chem. 270, 1395-1401, 1995  
 A;Title: A Xenopus nonmuscle myosin heavy chain isoform is phosphorylated by cyclin-p34 (Cyclin B)  
 A;Reference number: A55441; MUID:95138137; PMID:7836406  
 A;Accession: A55441  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 198-232 <REL>  
 C;Superfamily: myosin heavy chain; myosin motor domain homology  
 C;Keywords: nucleotide binding; P-loop  
 F;88-787/Domain: myosin motor domain homology <MMOT>  
 F;178-185/Region: nucleotide-binding motif A (P-loop)

Query Match 11.5%; Score 129.5; DB 2; Length 1992;  
 Best Local Similarity 26.0%; Pred. No. 3;  
 Matches 56; Conservative 44; Mismatches 76; Indels 39; Gaps 9;

QY 16 QVKKFSRAQKVLQKAVETKDERFEQASNFYQQAEGHKLKYLKXFLSAVKVM 74  
 DB 1331 QLQDTQELLQEBTRQKLNQ--SRIRQLKEEKNLQEQEERARKSLKQLSLQSQL 1388

QY 75 HESSKRVSEYLOIYSSWDGHBEEL-KAIWVNNDLLMEDYEK-----ADQA 121  
 DB 1389 TEAKKKVDDEVGTI-----EGLEEVKKLLKDTQGLQRLKEKIIAYEKLKTNLQOE 1443

QY 122 VRTMEIYVAFSEIKERIAKRGKLVYD-----SARHLE---AVONAKDEAKT 169  
 DB 1444 LDDLVDLDHQRQIVSNLEKKQK--FDQLLAEKNISARHAEERDRAEDAREKETKA 1500

QY 170 ---AKAEFEFNKAQTVFEDLNQELLELPILYNSR 201  
 DB 1501 LSLPALDEALEAQDEPERLNKQLRAEMEDLMSSK 1535

RESULT 11  
 S52537  
 emm L15 protein - Streptococcus pyogenes  
 C;Species: Streptococcus pyogenes  
 C;Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 28-May-1999  
 C;Accession: S52537  
 R;Katerov, V.; Schalen, C.; Totolian, A.A.  
 Mol. Gen. Genet. 245, 78-85, 1994  
 A;Title: Sequencing of genes within the vir regulon of Streptococcus pyogenes type M15  
 A;Reference number: S52535; MUID:95147851; PMID:7845360  
 A;Accession: S52537  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-377 <KAT>  
 A;Cross-references: GB:S75411; NID:g914107; PIDN:AAB33262.1; PID:g914110  
 C;Superfamily: M5 protein

```

Query Match      11.3%; Score 127; DB 2; Length 377;
Best Local Similarity 24.9%; Pred. No. 0.65; Mismatches 33; Indels 72; Gaps 8;
Matches 55; Conservative 33;

QY 19 KFSRAQEKVLQKLG-----KAVETKDRFPQSASNFYQQAEGHKLYKDLKFLSAVK 72
Db 68 EKLKENRYLEKIGQEBRQKNLEKQRFQVAADKHQYQVQKQHYKQ-----118
QY 73 VMHESKRYSETLOEITYSEWDGHEELKAIWNNDLLWEDYEELKLAQAVTMEIYVQAF 132
Db 119 -EGERKQKQOEQRKYQREVE-----KRYQEQLOKQ-----QQ 151
QY 133 SEIKERIAKRGKLVYDSARHLEAVONAKKD-EAKTAKAEEB-----175
Db 152 LETEKQISEARSKLSRD-----LEASRAAKKQLEAHEQKLEKEKQISDASRESSRDLE 206
QY 176 --FNKAQTVPEDL-----NOELLELPILYNSRIGCVVTI 208
Db 207 ASRESKSKVEADLAALTAHEQKLEKEKQISDASRQGLSVTL 247

RESULT 12
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Dictyostelium discoideum
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 19-Apr-2002
C;Accession: A26655; A24728; S00250
R;Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A;Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A;Reference number: A26655; MUID:87092266; PMID:3540939
A;Accession: A26655
A;Molecule type: DNA
A;Residues: 1-2116 <WAR>
A;Cross-references: GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1; PID:g167835
R;DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 6807-6810, 1995
A;Reference number: A24728; MUID:86016788; PMID:3901008
A;Accession: A24728
A;Molecule type: mRNA
A;Residues: 2035-2116 <DEL>
R;Waggle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A;Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium myosin heavy chain
A;Reference number: S00250; MUID:88112226; PMID:2828113
A;Accession: S00250
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1734-1893 <WAG>
C;Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphatase
F;1-818/Domain: globular head <HED>
F;89-747/Domain: myosin motor domain homology <MMOT>
F;179-186/Region: nucleotide-binding motif A (P-loop)
F;819-2116/Domain: alpha-helical rod <ROD>

Query Match      10.7%; Score 120.5; DB 2; Length 2116;
Best Local Similarity 22.7%; Pred. No. 12;
Matches 54; Conservative 39; Mismatches 90; Indels 55; Gaps 5;

QY 14 AKQVKQFSRAQEKVLQKLGAVETKDRFPQSASNFYQQAEGHKLYKDLKFLSAVKV 73
Db 1277 AKQALEKKRLGLESLKHNQLEBEKKQKESNEKRVLDLEKEVSELKQDIESEVASKKA 1336
QY 74 MHESKRYSETLOEIT-----YSEWDGHEELKAIWNNDLLWEDYEELKLAQAVTMEI 126
Db 1337 VTEAKNKKSELDLTKRYADVWSRDKSVLEQLATLQAKNEELRNRTAE-----1385
QY 127 IVVAFQFSIKERIAKRGKLVYDSARHLEAVONAKKDEAKTAKAEEFNKAQTVPED- 185

Query Match      10.5%; Score 118; DB 2; Length 399;
Best Local Similarity 23.3%; Pred. No. 2.5;
Matches 45; Conservative 43; Mismatches 81; Indels 24; Gaps 5;

QY 18 QKFSRAQEK-----VLQKLGAVETKDRFPQSASNFYQQAEGHKLYKDLKKNF 67
Db 94 EKKFAPVHEKKAESTAEKAAEVTEKVEIEELLEKEKSEKSELAKKVEIEHKVVEELEEK 153
QY 68 LSAVKVMHESKRYSETLOEITYSEWDGHEELKAIWNNDLLWEDYEELKLAQAVTMEI 127
Db 154 LTGEK-LEETKKYVEELEKIEKGEVTAEEVGELA-----EKVEELEAAKPEEE 204
QY 128 YVAFQFSIKERIAKRGKLVYDSARHLEAVONAKKDEAKTAKAEEFNK---AQTV 182
Db 205 VVVEVPEEEKVEVVEEVPPEEEVTPPEVEVEEKEEVEEKEEVEGKQSEMKKI 264
QY 183 FEDLNQELLLELP 195
Db 265 PEDIASILFPEEP 277

RESULT 14
A59287
myosin heavy chain - fluke (Schistosoma mansoni) (strain Brazilian LE)
C;Species: Schistosoma mansoni
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
C;Accession: A59287
R;Weston, D.S.; Schmitz, J.; Kemp, M.; Kunz, W.
Mol. Biochem. Parasitol. 58, 161-164, 1993
A;Title: Cloning and sequence characterization of a complete myosin heavy chain cDNA from Schistosoma mansoni
A;Reference number: A59287; MUID:93211444; PMID:8459827
A;Accession: A59287
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1940 <WES>
A;Cross-references: GB:L01634; PIDN:AAA29905.1
A;Experimental source: strain Brazilian LE
C;Genetics: MYH
A;Gene: MYH
C;Superfamily: myosin heavy chain; myosin motor domain homology
F;82-752/Domain: myosin motor domain homology <MMO>

Query Match      10.3%; Score 116; DB 2; Length 1940;
Best Local Similarity 26.2%; Pred. No. 20;
Matches 50; Conservative 29; Mismatches 76; Indels 36; Gaps 7;

```





Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	663.5	59.0	588	1	BIN1_MOUSE	O08539 mus musculus
2	663.5	59.0	588	1	BIN1_RAT	O08839 rattus norv
3	657.5	58.4	593	1	BIN1_HUMAN	O00459 homo sapien
4	620.5	55.2	683	1	AMPH_RAT	O08838 rattus norv
5	620.5	55.2	695	1	AMPH_HUMAN	P49418 homo sapien
6	618.5	55.0	682	1	AMPH_CHICK	P50478 gallus gall
7	192.5	17.1	264	1	HOB3_SCHPO	Q90477 schizosacch
8	179.5	16.0	265	1	R161_YEAST	P25343 saccharomyc
9	158.5	14.1	466	1	HOB1_SCHPO	O74352 schizosacch
10	154.5	13.7	482	1	R167_YEAST	P39743 saccharomyc
11	143	12.7	253	1	BIN3_HUMAN	Q9NqY0 homo sapien
12	122.5	10.9	1433	1	REST_CHICK	O24184 gallus gall
13	120.5	10.7	2116	1	MSX2_DICDI	R08799 dictyosteli
14	118.5	10.5	6885	1	SNR2_HUMAN	Q8wxh0 homo sapien
15	116	10.3	3259	1	G0B1_HUMAN	Q14789 homo sapien
16	114	10.1	226	1	Y637_PYRHO	O58571 pyrococcus
17	113	10.0	1005	1	R450_METUA	O58718 methanococc
18	113	10.0	1300	1	DYNA_NEUCR	Q001397 neurospora
19	113	10.0	1972	1	MYHB_MOUSE	O08638 mus musculus
20	112.5	9.9	2245	1	MSX3_DICDI	P54697 dictyosteli
21	111.5	9.9	1790	1	USO1_YEAST	P25386 saccharomyc
22	110.5	9.8	1169	1	SWC_METUA	O59037 methanococc
23	110	9.8	288	1	Y187_CABEL	P34445 caenorhabdi
24	110	9.8	407	1	M21_STRPY	P50468 streptococc
25	110	9.8	1972	1	MYHB_HUMAN	P35749 homo sapien
26	110	9.8	1976	1	MYHA_RAT	Q9J1t0 rattus norv
27	109.5	9.7	944	1	NUF1_YEAST	P33380 saccharomyc
28	109.5	9.7	1225	1	SMC1_YEAST	P33908 saccharomyc
29	109	9.7	284	1	TPM_ECHMU	Q95pul echinococcu
30	109	9.7	520	1	REC1_AQUAB	O68834 aquifex aeo
31	109	9.7	942	1	SYL_METUA	Q58050 methanococc
32	109	9.7	1170	1	SMC2_YEAST	P39898 saccharomyc
33	109	9.7	2017	1	MYSN_DRONE	Q93923 drosophila

```

CC -!- TISSUE SPECIFICITY: Isoform 1 is expressed mainly in the
CC brain. Isoform 2 is widely expressed.
CC -!- PTM: Phosphorylated by protein kinase C (By similarity).
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U86405; AAC53318.1; -
CC EMBL; U60884; AAC52661.1; -
CC MGD; MG:108092; Bin1.
CC InterPro; IPR003005; Amphiphysin.
CC InterPro; IPR006632; BAR.
CC InterPro; IPR004148; BAR dom.
CC InterPro; IPR001452; SH3.
CC Pfam; PF03114; BAR; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR01251; AMPHIPHYSIN.
CC PRINTS; PR0452; SH3DOMAIN.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00721; BAR; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
CC KW Alternative splicing; SH3 domain; Coiled coil; Endocytosis;
CC Anti-oncogene; Differentiation; Phosphorylation.
CC FT DOMAIN 15 42
CC FT COILED COIL (POTENTIAL).
CC FT DOMAIN 193 274
CC FT COILED COIL (POTENTIAL).
CC FT DOMAIN 379 422
CC FT CLATHRIN-BINDING (BY SIMILARITY).
CC FT SH3.
CC FT VARSPLIC 174 204
CC FT Missing (in isoform 2).
CC FT VARSPLIC 335 457
CC FT Missing (in isoform 2).
CC FT VARSPLIC 000255.
CC FT VARSPLIC 000256.
CC SQ SEQUENCE 588 AA; 64470 MW; 63CA362461500F38 CRC64;

Query Match 59.0%; Score 663.5; DB 1; Length 588;
Best Local Similarity 54.3%; Pred No. 9.3e-33;
Matches 138; Conservative 30; Mismatches 53; Indels 33; Gaps 3;

QY 1 MAE-GKAGAGLFAKQVKFSAQKVLQKLGKAVETKDRPEQBSANFYQQAGHGK 59
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 60 LYKDLKNFLSAVKVHESKRVSETLQELYSSEWDGHEELKAIWVNDLLWEDYEKLD 119
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 LQDRTYLASVKAHESKSLCEQBSVPEPWPGRDEANKIAENNLLMDYHOKLVD 120
QY 120 QAVRTMEIYVAQFSEIKERIAKRGKLVYDYSARHLEAVQNA-KDEAKTAK----- 171
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 121 QALLTMDTVLGFPDPIKSRIAKRGKLVYDYSARHLYESLTACKDEAKIAKVSLEK 180
QY 172 -----AEEFNKAQTVFEDLNQELLEHELPILYNRGICVYT 207
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 181 AAPQWQQKQLAHVAQTNLLENQAEEELIKAKVFPEMNVYDLOEELPSLWNSRVGFYN 240
QY 208 IFQINISNLRDVFYR 221
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 241 TFQSIAGLEENPHK 254

RESULT 2
ID BIN1.RAT
AC O08839;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)

```

```

DE DE -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).
GN BIN1 OR AMPH1 OR AMPH2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.; SUBUNIT, AND ALTERNATIVE SPLICING.
RC STRAIN=Sprague-Dawley; TISSUE=Brain cortex, and Kidney;
RX MEDLINE=98009145; PubMed=9348539;
RA Wigge P., Kohler K., Hunt S., Doyle C., McMahon H.T.;
RT "Amphiphysin heterodimers: potential role in clathrin-mediated
RT endocytosis".
RL Mol. Biol. Cell 8:2003-2015 (1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM AMPH2-1).
RC STRAIN=Sprague-Dawley; TISSUE=Brain cortex;
RX MEDLINE=97424383; PubMed=9280305;
RA McMahon H.T., Wigge P., Smith C.;
RT "Clathrin interacts specifically with amphiphysin and is displaced by
RT dynamin.".
RL FEBS Lett. 413:319-322 (1997).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 506-588, FUNCTION, AND
RP INTERACTION WITH DNM1.
RX MEDLINE=98409533; PubMed=9736607;
RA Owen D.J., Wigge P., Vallis Y., Moore J.D.A., Evans P.R.,
RA McMahon H.T.;
RT "Crystal structure of the amphiphysin-2 SH3 domain and its role in the
RT prevention of dynamin ring formation.".
RL EMBO J. 17:5273-5285 (1998).
CC -!- FUNCTION: May be involved in regulation of synaptic vesicle
CC endocytosis. May act as a tumor suppressor and inhibits
CC malignant cell transformation.
CC -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).
CC Binds to SYN1 and DNM1 through its SH3 domain, and to clathrin
CC through a region outside of the SH3 domain. Also binds to alpha-
CC adaptin. Interacts with the N-terminal transactivation domain of
CC MYC in a manner requiring the integrity of the conserved MYC box
CC regions 1 and 2.
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=AMPH2-1;
CC IsoId=O08839-1; Sequence=Displayed;
CC Name=AMPH2-2;
CC IsoId=O08839-2; Sequence=VSP_000260;
CC Name=AMPH2-3;
CC IsoId=O08839-3; Sequence=VSP_000258;
CC Name=AMPH2-4;
CC IsoId=O08839-4; Sequence=VSP_000257;
CC Name=AMPH2-5;
CC IsoId=O08839-5; Sequence=VSP_000259;
CC Name=AMPH2-6;
CC IsoId=O08839-6; Sequence=VSP_000256, VSP_000259;
CC -!- TISSUE SPECIFICITY: Isoform AMPH2-1 is expressed in brain,
CC concentrated at nerve terminals. Isoform AMPH2-2 is widely
CC expressed.
CC -!- PTM: Phosphorylated by protein kinase C.
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y13380; CAA73807.1; -
CC PDB; 1BB9; 17-JUN-98.

```

DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR dom.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW Alternative splicing; SH3 domain; Coiled coil; Endocytosis;  
 FT Anti-oncogene; Differentiation; Phosphorylation; 3D-structure.  
 FT DOMAIN 15 42  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 193 274  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 379 422  
 FT CLATHRIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 515 588  
 FT SH3.  
 FT VARSPLIC 173 205  
 FT Missing (in isoform AMPH2-4 and isoform AMPH2-6).  
 FT FTID=VSP 000256.  
 FT Missing (in isoform AMPH2-4).  
 FT FTID=VSP 000257.  
 FT Missing (in isoform AMPH2-3).  
 FT FTID=VSP 000258.  
 FT Missing (in isoform AMPH2-5 and isoform AMPH2-6).  
 FT FTID=VSP 000259.  
 FT Missing (in isoform AMPH2-2).  
 FT FTID=VSP 000260.  
 FT SEQUENCE 588 AA; 64533 MW; 164AC90E9547F1A CRC64;  
 Query Match 59.0%; Score 663.5; DB 1; Length 588;  
 Best Local Similarity 54.3%; Pred. No. 9.3e-33;  
 Matches 138; Conservative 30; Mismatches 53; Indels 33; Gaps 3;  
 QY 1 MAE-GKAGAGLPAKQVOKFSPRAQKVLQKAVETKDERPEQASFYQQAEHGH 59  
 Db 1 NAEAGSGKVTAGKIASNVQKLTAAQKVLQKAGADTKDEQECVQVFNKQLTEGR 60  
 QY 60 LYKDLNPLSAVKVHSESKVSETLQEIYSSEWDGHEELKAIWNNDLLWEDYEKLAD 119  
 Db 61 LQDLRTYLVASKVHSESKVSECLQVPEPEWFGDEANKAENNDLLWMDYHOKLVD 120  
 QY 120 QAVRTMEIYVQSFSEIKRIARQKGLVDYDSARHHLVAVQNA-KKDEAKTAK----- 171  
 Db 121 QALLTMDTYLGQFPDIKSRIAKRGKLVYDSARHLYVESLQAKKDEAKIAPVSLLEK 180  
 QY 172 -----AEEFNKAQVTFEDLNQELLELPLIYNRIGCYVT 207  
 Db 181 AAPQWQCKQLQHLVQATNLRNQAEEELIKAKVFEEMNVDLQELFSLWNSRVGFYN 240  
 QY 208 IFQINSLNLDVYFR 221  
 Db 241 TFSIAGLEENFHK 254  
 RESULT 3  
 ID BIN1 HUMAN STANDARD; PRT; 593 AA.  
 AC O00499; O00297; O00545; O43867; O60552; O60553; O60554; O60555;  
 AC O75514; O75515; O75516; O75517; O75518; O92944; Q99688;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Myc box dependent interacting protein 1 (Bridging integrator 1)  
 DE (Amphiphysin-like protein) (Amphiphysin II) (Box-dependent myc-interacting protein-1).  
 GN BIN1 OR AMPHL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM IIA).  
 RC TISSUE=Brain;  
 RX MEDLINE=97341217; PubMed=9195986;  
 RA Ramjaun A.R., Micheva K.D., Bouchelet I., McPherson P.S.;  
 RT "Identification and characterization of a nerve terminal-enriched amphiphysin isoform";  
 RL J. Biol. Chem. 272:16700-16706 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS IIA AND BIN1).  
 RC TISSUE=Brain, and skeletal muscle;  
 RX MEDLINE=97327761; PubMed=9182667;  
 RA Butler M.H., David C., Ochoa G.-C., Freyberg Z., Daniell L., Grabs D., Crenona O., De Camilli P.;  
 RT "Amphiphysin II (SH3P9; BIN1), a member of the amphiphysin/Rvs family, is concentrated in the cortical cytomatrix of axon initial segments and nodes of ranvier in brain and around T tubules in skeletal muscle";  
 RL J. Cell Biol. 137:1355-1367 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM BIN1).  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=96376973; PubMed=8782822;  
 RA Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
 RT "BIN1 is a novel Myc-interacting protein with features of a tumour suppressor";  
 RL Nat. Genet. 14:69-76 (1996).  
 RN [4]  
 RP REVISIONS TO N-TERMINUS.  
 RA Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS IIB; IIC1; IIC2 AND IID).  
 RC TISSUE=Brain;  
 RX MEDLINE=98264340; PubMed=9603201;  
 RA Ramjaun A.R., McPherson P.S.;  
 RT "Multiple amphiphysin II splice variants display differential clathrin binding: identification of two distinct clathrin-binding sites";  
 RL J. Neurochem. 70:2369-2376 (1998).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS I12 AND I13).  
 RC TISSUE=Brain;  
 RX MEDLINE=97366618; PubMed=9223448;  
 RA Tsutsui K., Maeda Y., Tsutsui K., Seki S., Tokunaga A.;  
 RT "cDNA cloning of a novel amphiphysin isoform and tissue-specific expression of its multiple splice variants";  
 RL Biochem. Biophys. Res. Commun. 236:178-183 (1997).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORMS I13; I13; BIN1-10-13; BIN1-13 AND BIN1-12A).  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=98058932; PubMed=9395479;  
 RA Wechsler-Reya R.J., Sakamuro D., Zhang J., DuHadaway J., Prendergast G.C.;  
 RT "Structural analysis of the human BIN1 gene. Evidence for tissue-specific transcriptional regulation and alternate RNA splicing";  
 RL J. Biol. Chem. 272:31453-31458 (1997).  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORM I12).  
 RA Zhang J., Du W., Wechsler-Reya R.J., DuHadaway J., Sakamuro D., Prendergast G.C.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE OF 133-593 FROM N.A.  
 RC TISSUE=Brain;  
 RA Yu W., Gibbs R.A.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [10]  
 RP CHARACTERIZATION.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=98078712; PubMed=9418903;  
 RA Wechsler-Reya R.J., Elliott K.J., Prendergast G.C.;  
 RT "A role for the putative tumor suppressor Bin1 in muscle cell



```

DE Amphiphsin.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain cortex;
RX MEDLINE=98009145; PubMed=9348539;
RA Wiggie P., Kohler K., Vallis Y., Owen D., Hunt S.P., McMahon H.T.;
RT "Amphiphsin heterodimers: potential role in clathrin-mediated
RL endocytosis.";
RL Mol. Biol. Cell 8:2003-2015(1997).
CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in
CC synapses and certain endocrine cell types. May control the
CC properties of the membrane associated cytoskeleton (By
CC similarity).
CC -!- SUBUNIT: Heterodimer with BIN1. Binds SH3GLB1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Peripheral membrane protein associated with
CC the cytoplasmic surface of synaptic vesicles (By similarity).
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DE BMBL; Y13381; CAA73808.1;
DR InterPro; IPR003005; Amphiphsin.
DR InterPro; IPR003017; Amphiphsin_1.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR dom.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PRO1251; AMPHIPHYSIN.
DR PRINTS; PRO0452; SH3DOMAIN.
DR ProDom; PD003208; Amphiphsin_1; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Cytoskeleton; SH3 domain; Coiled coil.
FT DOMAIN 10 83 COILED COIL (POTENTIAL).
FT DOMAIN 144 191 COILED COIL (POTENTIAL).
FT DOMAIN 610 683 SH3.
SQ SEQUENCE 683 AA; 74877 MW; 7FEA4A9E5A1F6631 CRC64;

Query Match
Best Local Similarity 55.2%; Score 620.5; DB 1; Length 683;
Matches 116; Conservative 49; Mismatches 46; Indels 1; Gaps 1;

QY 11 GLFAQVQKFSRAQEKVQLGKAVETKDERFEQASNFYQQAGHKLYKDLNFLSA 70
Db 7 GIFAQVQKLNRAQEKVQLGKADETKDEFEYVQFKQAEAGTQLQELRGYLA 66
QY 71 VYVHESKRVSELTQBIYSSEWDGHEELKAIYVNNLLMEDYEEKLADQAVRTWEIYVA 130
Db 67 IKGQASMKLTESLHEVYFPDWMYGRDVKVNGEKDVLWEDPHQKLVGSLTLDYLG 126
QY 131 QFSIKRIAKRGKLVYDSARHLEAVQNAK-KDEAKTAKAEFPNKAQVTFEDLNQE 189
Db 127 QFPDKRIAKRGKLVYDSARHLEALQSSRRKDESRISKAEFFQKAQKVFPEFND 186
QY 190 LLEELPILYNSRGYVTFIQNISNLRDVFYR 221
Db 187 LQELPSLWRRGVFYNTFNKVSLEAKFKH 218

```

```

RESULT 5
AMPH_HUMAN
ID ID AMPH_HUMAN STANDARD; PRT; 695 AA.
AC P49418; O43538;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amphiphsin.
GN AMPH OR AMPH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX MEDLINE=94357284; PubMed=8076697;
RA David C., Solimena M., de Camilli P.;
RT "Autoimmunity in stiff-man syndrome with breast cancer is targeted to
RT the C-terminal region of human amphiphsin, a protein similar to the
RT yeast proteins, Rvs167 and Rvs161.";
RL FEBS Lett. 351:73-79(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=95276740; PubMed=7757077;
RA Yamamoto R., Li X., Winter S., Francke U., Kilian M.W.;
RT "Primary structure of human amphiphsin, the dominant autoantigen of
RT paraneoplastic stiff-man syndrome, and mapping of its gene (AMPH) to
RT chromosome 7p13-p14.";
RL Hum. Mol. Genet. 4:265-268(1995).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Breast;
RX MEDLINE=98174372; PubMed=9513187;
RA Floyd S.R., Butler M.H., Cremona O., David C., Freyberg Z., Zhang X.,
RA Solimena M., Tokunaga A., Ishizu H., Tsutsui K., De Camilli P.V.;
RT "Expression of amphiphsin I, an autoantigen of paraneoplastic
RT neurological syndromes, in breast cancer.";
RL Mol. Med. 4:29-39(1998).
CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in
CC synapses and certain endocrine cell types. May control the
CC properties of the membrane associated cytoskeleton.
CC -!- SUBUNIT: Heterodimer of AMPH and AMPH1 (By similarity).
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC SURFACE OF
CC SYNAPTIC VESICLES.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=128 kDa;
CC IsoId=P49418-1; Sequence=Displayed;
CC Name=2; Synonyms=108 kDa;
CC IsoId=P49418-2; Sequence=VSP_000245;
CC -!- TISSUE SPECIFICITY: Neurons, certain endocrine cell types and
CC spermatocytes.
CC -!- DISEASE: Patients with stiff-man syndrome, a rare disease of the
CC central nervous system characterized by progressive rigidity of
CC the body musculature with superimposed painful spasms, have
CC antibodies against AMPH.
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DE BMBL; U07616; AAA21865.1;
DR DR EMBL; X81438; CAA57197.1;
DR EMBL; AF034996; AAC02977.1;

```

[illegible]







Db 64 LNHQGFADACIEIYKPIGRASDPESVEQGNAGIEBAEAYKEIYV-----DLQKN 116  
 QY 117 LADQAVRTMEIYVAF-----SEIKERIAKRGKLVDSARHLEAVONAK-- 163  
 Db 117 LASE-----MDVINTRIVNPTGELLKIVKDVDDKLLKRDHKQLDYDRHRSFKLQEKQK 172  
 QY 164 --KDEAKTAAEEFNKAQTVFEDLNQELBELFPIYNSRIGCVTVTFQINSLNR-DVFFY 220  
 Db 173 SLKDEKKIYEATAFEQSSQVEYYNEMLKEELPKLPALAQSFIAFLFQGYVQNLVYY 232

RESULT 10  
 R167\_YEAST  
 ID R167\_YEAST STANDARD; PRT; 482 AA.  
 AC P39743;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Reduced viability upon starvation protein 167.  
 GN RVS167 OR YDR388W OR D9509.8.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180;  
 RX MEDLINE=93330299; PubMed=8336735;  
 RA Bauer F., Urdaci M., Aigle M., Crouzet M.;  
 RT "Alteration of a Yeast SH3 protein leads to conditional viability  
 with defects in cytoskeletal and budding patterns.";  
 RL Mol. Cell. Biol. 13:5070-5084(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97313263; PubMed=9169867;  
 RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Bahr A., Ballesta J.P.G.,  
 RA Barques M., Baron L., Becker N., Bloeker H., Bugnon C.,  
 RA Boskovic J., Brandt P., Bruckner M., Buitrago M.J., Coster F.,  
 RA Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,  
 RA Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Hankeln T.,  
 RA Hohnel J.D., Jaeger W., Jimenez A., Joniaux J.-L., Kraemer C.,  
 RA Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,  
 RA Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,  
 RA Paulin L., Perea J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,  
 RA Prydz H., Purnelle B., Rasmussen S.W., Remacha M., Revuelta J.B.,  
 RA Rieger M., Salom D., Saluz H.P., Salz J.E., Saren A.-M., Schaefer M.,  
 RA Scharfe M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,  
 RA Urrastazu L.A., Verhaesselt P., Vissers S., Voet M., Volckaert G.,  
 RA Wagner G., Wambutt R., Wedler E., Wedler H., Wolfel S., Harris D.E.,  
 RA Bowman S., Brown D., Churcher C.M., Connor R., Dedman K., Gentles S.,  
 RA Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D.,  
 RA Odell C., Oliver K., Rajandream M.A., Richards C., Shore L.,  
 RA Walsh S.V., Barrall B.G., Dietrich F.S., Mulligan J.T., Allen E.,  
 RA Araujo R., Aviles E., Berno A., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung B., Duncan M., Hunnicke-Smith S., Hyman R.W., Komp C.,  
 RA Laskhari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A.,  
 RA Oefner P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M.,  
 RA Shogren T., Shroff N., Winant A., Yelton M.A., Botstein D.,  
 RA Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,  
 RA Du Z., Favell A., Fulton L., Gattung S., Greco T., Hallsworth K.,  
 RA Hawkins J., Hillier L.W., Jier M., Johnson D., Johnston L.,  
 RA Kirsten J., Kucaba T., Langston Y., Latreille P., Le T., Mardis E.,  
 RA Menezes S., Miller N., Nhan M., Pauley A., Peluso D., Rifkin L.,  
 RA Riles L., Taich A., Trevasakis E., Vignati D., Wilcox L., Wohldman P.,  
 RA Vaudin M., Wilson R., Waterston R., Albermann K., Hani J., Heumann K.,  
 RA Kleine K., Mewes H.-W., Zollner A., Zaccaria P.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";  
 RL Nature 387:75-78 (1997).  
 RN [3]  
 RP ACTIN-BINDING.  
 RX MEDLINE=95236199; PubMed=7719850;  
 RA Amberg D.C., Basart E., Botstein D.;  
 RT "Defining protein interactions with yeast actin in vivo.";

RL Nat. Struct. Biol. 2:28-35 (1995).  
 CC -!- FUNCTION: Component of a cytoskeletal structure that is required  
 CC for the formation of endocytic vesicles at the plasma membrane  
 CC level. Could be implicated in cytoskeletal reorganization in  
 CC response to environmental stresses and could act in the budding  
 CC site selection mechanism. Binds to actin.  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M92092; AAA35051.1; -;  
 CC EMBL; U32274; AAB64830.1; -;  
 CC PIR; S40887; S40887.  
 CC HSSP; P19174; 2HSP.  
 CC GerMOnline; 140880; -;  
 CC SGD; S0002796; RVS167.  
 CC GO; GO:0005857; C:actin cortical patch (sensu Saccharomycetes); IDA.  
 CC GO; GO:0008092; F:cytoskeletal protein binding; IPI.  
 CC GO; GO:0006897; P:endocytosis; IMP.  
 CC GO; GO:0007121; P:polar budding; IMP.  
 CC GO; GO:0006970; P:response to osmotic stress; IMP.  
 CC InterPro; IPR0006532; BAR.  
 CC InterPro; IPR004148; BAR\_dom.  
 CC InterPro; IPR001452; SH3.  
 CC Pfam; PF03114; BAR; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC PRINTS; PR00452; SH3DOMAIN.  
 CC ProDom; PD000066; SH3; 1.  
 CC SMART; SM00721; BAR; 1.  
 CC SMART; SM00326; SH3; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 CC Cytoskeleton; SH3 domain; Coiled coil; Transmembrane; Actin-binding.  
 FT DOMAIN 31 64  
 FT DOMAIN 174 204  
 FT DOMAIN 292 427  
 FT TRANSMEM 344 367  
 FT DOMAIN 421 482  
 FT SEQUENCE 482 AA; 52774 MW; 3F0AB53EBCC95A5B CRC64;  
 SQ  
 Query Match 13.7%; Score 154.5; DB 1; Length 482;  
 Best Local Similarity 24.4%; Pred. No. 0.015;  
 Matches 61; Conservative 38; Mismatches 100; Indels 51; Gaps 8;  
 QY 15 KQVKFSRAQEKVLQKLGAVETKDERFQSSNFYQQAEGHKLKYDKLNFLSAV--- 71  
 Db 4 KGFTKVASRAQSFQKFRKMGEGTDPVVEDARRRFOELEQETKQLSSESKRYSTAVNGM 63  
 QY 72 -----KWHSSSKRVS-----ETIQEIVSSEWDGHEELKAVVN-NDLLMED--- 112  
 Db 64 LTHQIGFAKSENEIFKPIGKMSDPNATIPENPQGEASEQYRAIVAELOETLPDLAL 123  
 QY 113 YEEKLADQAVRTMEIYVAFSEIKERIAKRGKLVDSARHLEAVONAKDEAKTAKA 172  
 Db 124 VEEKIVTPCQELKI-----ITVIRKMATKGNKHKLLDLD---RLHNTYKHKHKKKPTAKD 176  
 QY 173 EESFNKAQTVFEDLNQ-----LLEELPIL-----YNSRIGCVTVTFQ 210  
 Db 177 EERLYKAAQAVQVAAQOEYDYNDLNTQLTPILFSLAEAFVKPLFVSVFYFMQLNIFITLYN 236  
 QY 211 NISNLRDVFFY 220  
 Db 237 RLQDMKIPYF 246

RESULT 11  
 BIN3\_HUMAN

```

DR Pfam; PF03114; BAR; 1.
DR SMART; SM00721; BAR; 1.
DR Septation; Cytoskeleton; Coiled coil.
DR FT DOMAIN 12 208 BAR.
FT FT DOMAIN 18 51 COILED COIL (POTENTIAL).
FT FT DOMAIN 120 152 COILED COIL (POTENTIAL).
FT FT DOMAIN 231 247 COILED COIL (POTENTIAL).
SQ SEQUENCE 253 AA; 29665 MW; 254CC7113749C584 CRC64;

Query Match 12.7%; Score 143; DB 1; Length 253;
Best local similarity 22.7%; Pred. No. 0.037;
Matches 48; Conservative 49; Mismatches 18; Gaps 5;

QY 20 KFSRAQKVLQKLGKAVETKDFEFOSAGNPFQQQAEGHKLYKDLGNPLSAVKVHSSK 79
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
7 KIGQPKQIVK-----TVRDFEREYKGLQQLSEQTRRLQKMKKSTDLAMSKAV 60
QY 80 RVSELTQIYSSWDGHEELKATVWNNDLL---WEDYEKLDAQVRT---MEIYVAQF 132
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
61 KISLDLNSPLCEQD--QDLLNMTALDTAMKRMDFAPNQKVNQIQKTIVIEPLKKGVSF 118
QY 133 SEIKERIAKGRKLVYDVSARHLEAVQNAKDE---AKTAKASEEFNKAQTVFEDLNQE 189
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
119 PSLNAVVRREQALQDYRLQAKVEKYBEKENTGPVLAKLHQARELFPVRDEPEAKNQ 178
QY 190 LLEELPILYNSRICGYVTIFONISNLRDVFY 220
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
179 LLEENPRYGRLDYFQPSFESLIRAQVVIY 209

RESULT 12
REST CHICK
ID REST CHICK STANDARD; PRET; 1433 AA.
AC Q42184; Q42228; Q57563; Q57564;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Restin (Cytoplasmic linker protein-170) (CLIP-170).
DE RSN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
[1]
RN MEDLINE=98137792; PubMed=9469933;
RX GRIPARIC L., Volosky J.M., Keller T.C. III;
RA "Cloning and expression of chicken CLIP-170 and restin isoforms.";
RT Gene 206:195-208(1998).
RN [2]
RN SEQUENCE OF 17-1139 FROM N.A. (ISOFORMS 3 AND 4).
RP TISSUE=Pectoralis muscle;
RC MEDLINE=93002898; PubMed=9784600;
RA GRIPARIC L., Keller T.C. III;
RT "Identification and expression of two novel CLIP-170/Restin isoforms expressed predominantly in muscle.";
RL Biochim. Biophys. Acta 1405:35-46(1998).
CC -1- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE CYTOSKELETON (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q42184-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q42184-2; Sequence=VSP_000761;
CC Name=3; Synonyms=CLIP-170(11);
CC IsoId=Q42184-3; Sequence=VSP_000762, VSP_000763;
CC Name=4; Synonyms=CLIP-170(11+35);
CC IsoId=Q42184-4; Sequence=VSP_000764;

```

```

CC -1- SIMILARITY: Contains 2 CAP-Gly domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ihb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL: AF014012; AAC60344.1; -
DR EMBL: AF020764; AAC60345.1; -
DR EMBL: AF045650; AAC03547.1; -
DR EMBL: AF045651; AAC03548.1; -
DR InterPro: IPR000938; CAP-Gly.
DR InterPro: IPR001878; Znf CCHC.
DR Pfam: PF01302; CAP GLY; 2.
DR SMART: SM00343; Znf C2HC; 1.
DR PROSITE: PS00845; CAP_GLY_1; 2.
DR PROSITE: PS02445; CAP_GLY_2; 2.
KW Cytoskeleton; Microtubule; Coiled coil; Repeat; Alternative splicing.
FT DOMAIN 79 121
FT DOMAIN 144 207
FT DOMAIN 235 277
FT DOMAIN 305 332
FT DOMAIN 351 1353
FT DOMAIN 1414 1427
FT VARSPPLIC 458 492
FT VARSPLIC 458 492
FT VARSPLIC 803 803
FT VARSPLIC 458 458
FT VARSPLIC 309 309
FT CONFLICT 440 440
FT CONFLICT 440 440
FT SEQUENCE 1433 AA; 161026 MW; 5631CE8683498E23 CRC64;
Query Match 10.9%; Score 122.5; DB 1; Length 1433;
Best Local Similarity 27.1%; Pred. No. 3.6;
Matches 58; Conservative 37; Mismatches 74; Indels 45; Gaps 10;
QY 15 KQVKKFSRAQKVLQKLGKAVETKDFEQAASNFVQQQAQGH-KLYKDLKNFLSAVKV 73
Db 943 EQIQLELTKANEKAVQ-LQKNVBTQAQAEQS-----QQETLKTQHEELKQMDQLTDMKK 997
QY 74 MHESSKRYSETLQEIYSS-----WDGHEELKAIWNNDLLWEDYE 114
Db 998 QWETSQVQKDLQAKYKESKEMTKHDADIKGPKQNLDAEALKAQKQNDLETOAE 1057
QY 115 E-----KLADQAVRTMEIYVQAQSEI-KERIAKGRKLVVDYDSARHHEAVQNA 162
Db 1058 ELKQAEQAKADKAEAEVLTQMEKVTKEKDAIHQEKI-----ETLASLENSRQTNKQLN- 1112
QY 163 KQDEAK--TAXAEFEFNKAQVFDLNOELLEEL 194
Db 1113 ELDMLKQNNLQNEBELTKSKELL-NLENKKVEEL 1145
RESULT 13
ID_MYS2_DICDI STANDARD; PRT; 2116 AA.
AC P08739;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dictyostelium II heavy chain, non muscle.
GN MHCA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;

```

```

RN RP
RX MEDLINE=87092266; PubMed=3540939;
RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;
RT "Conserved protein domains in a myosin heavy chain gene from
RT Dictyostelium discoideum.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN RP
RX MEDLINE=90353583; PubMed=2387408;
RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
RA Gerisch G.;
RT "Replacement of threonine residues by serine and alanine in a
RT phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL FEBS Lett. 269:239-243(1990).
RN RP
RX MEDLINE=88112226; PubMed=2828113;
RA Wagle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
RT Dictyostelium myosin heavy chain.";
RL FEBS Lett. 227:71-75(1988).
RN RP
RX MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RA Rayment I.;
RT "X-ray structures of the myosin motor domain of Dictyostelium
RT discoideum complexed with MgADP.BefX and MgADP.AIF4-";
RL Biochemistry 34:8960-8972(1995).
RN RP
RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT truncated head of Dictyostelium discoideum myosin to 2.7-A
RT resolution.";
RL Biochemistry 34:8973-8981(1995).
RN RP
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RT Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL Biochemistry 35:5404-5417(1996).
RN RP
RX MEDLINE=97452580; PubMed=9305951;
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammaS, and MgAMPNP complexes
RT of the Dictyostelium discoideum myosin motor domain.";
RL Biochemistry 36:11619-11628(1997).
RN RP
RX MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RT of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT Dictyostelium discoideum myosin motor domain.";
RL J. Mol. Biol. 274:394-407(1997).
CC -1- FUNCTION: Myosin is a protein that binds to actin and has ATPase
CC activity that is activated by actin.
CC -1- SUBUNIT: Myosin II heavy chain is two-headed. It self-assembles
CC into filaments. Hexamer of 2 heavy chain subunits (MHC), 2 alkali
CC light chain subunits (MLC) and 2 regulatory light chain subunits
CC (MLC-2).
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC CORTEX.
CC -1- DOMAIN: Each myosin heavy chain can be split into 1 light
CC meromyosin (LMM) and 1 heavy meromyosin (HMM). It can be further
CC split into 2 globular subfragments (S1) and 1 rod-shaped
CC subfragment (S2).

```

-1- DOMAIN: The rodlike tail sequence is highly repetitive, showing cycles of a 28-residue repeat pattern composed of 4 heptapeptides, characteristic for alpha-helical coiled coils.  
-1- PM: Phosphorylation inhibits thick filament formation and reduces the actin-activated ATPase activity.  
-1- MISCELLANEOUS: Dictyostellum myosin II has no K(2)EDTA ATPase activity, perhaps correlated with the absence of a Cys at the SH-1 position (688).  
-1- SIMILARITY: Contains 1 myosin-like globular head domain.  
-1- SIMILARITY: Contains 1 IQ domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

FT	STRAND	69	73
FT	HELIX	74	76
FT	STRAND	78	79
FT	HELIX	83	85
FT	TURN	86	87
FT	STRAND	90	90
FT	HELIX	91	93
FT	HELIX	99	110
FT	TURN	111	113
FT	STRAND	116	119
FT	TURN	120	121
FT	STRAND	122	126
FT	HELIX	137	142
FT	TURN	143	145
FT	HELIX	148	150
FT	HELIX	155	169
FT	TURN	170	170
FT	STRAND	173	179
FT	TURN	181	182
FT	HELIX	185	200
FT	HELIX	210	226
FT	STRAND	227	228
FT	TURN	234	235
FT	STRAND	236	237
FT	STRAND	240	247
FT	TURN	249	250
FT	STRAND	253	261
FT	HELIX	265	268
FT	TURN	269	269
FT	TURN	273	274
FT	STRAND	278	278
FT	HELIX	279	287
FT	HELIX	290	296
FT	TURN	297	297
FT	HELIX	301	303
FT	TURN	305	307
FT	TURN	316	317
FT	HELIX	320	334
FT	TURN	335	335
FT	HELIX	338	338
FT	TURN	356	356
FT	STRAND	360	360
FT	STRAND	368	368
FT	HELIX	373	382

Query Match	10.7%;	Score 120.5;	DB 1;	Length 2116;
Best Local Similarity	22.7%;	Pred. No. 7;		
Matches 54;	Conservative 39;	Mismatches 90;	Indels 55;	Gaps 5;

QY	14	AKQVQKFSBAQEKVLOKLGKAVTKDFERFQOSANFYQQOAEHGKLYKOLKNFLSAVKV	73
Db	1277	AKQALEKKRLGLESSELKHVNEQLEBEKKQKNEKKRVDLKEVSELDQOIBEEVASKKA	1336
QY	74	MHESKEVSTIQEI-----YSEWDGHELKAIWNNDLLWEDYEKLAQAVRTWE	126
Db	1337	VTEAKNCKESELDEIKROYADVSSRDKSVQQLTQAKQIEELRNTAE-----	1385
QY	127	IYVAQFSEIKERTAKGRKLVVDYSARHLEAVONAKDKDAKTAKABEEFNKAQTVFED-	185
Db	1386	-----AEGQLDRAERSKKAEFDLE----FAVNLESEETAKVKAEKAMKKAETDYRST	1435
QY	186	-----LNOEL-----LEELPILYNSRIGCVYTFIQNISLRD	217
Db	1436	KSELDDAKNVSSPOYVOIKRLNEELSELRSVLEBADERCNSAKKAKTASALESLKD	1493

RESULT 14	
SNE2 HUMAN	
ID	SNE2 HUMAN
STANDARD;	PRT; 6885 AA.
Q8WVX0; Q8N1S3;	Q8NFA9; Q8TER7; Q8WVW3; Q8WTW4; Q8WVW5; Q8WVX1;
Q8WVX5; Q8UFQ4;	Q8Y2L4; Q8Y4R1;
DT 10-OCT-2003	(Rel. 42, Created)
DT 10-OCT-2003	(Rel. 42, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)  
 DT Nespriin 2 (Nuclear envelope spectrin repeat protein 2) (Synne-2)  
 DE (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting  
 DE element protein) (NUANCE protein).  
 GN SYNE2 OR NUA OR KIAA0111.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND  
 RP INTERACTION WITH F-ACTIN.  
 RX MEDLINE=22113122; PubMed=12118075;  
 RA Zhen Y.-Y., Libotte T., Munk M., Noegel A.A., Korenbaum E.;  
 RT "NUANCE, a giant protein connecting the nucleus and actin  
 RT cytoskeleton.";  
 RL J. Cell Sci. 115:3207-3222 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 4; 5 AND 7).  
 RX MEDLINE=21652858; PubMed=11792814;  
 RA Zhang Q., Skepper J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G.,  
 RA Weisberg P.L., Ellis J.A., Shanahan C.M.;  
 RT "Nesprins: a novel family of spectrin-repeat-containing proteins that  
 RT localize to the nuclear membrane in multiple tissues.";  
 RL J Cell Sci. 114:4485-4498 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=22296983; PubMed=12408964;  
 RA Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;  
 RT "The nesprins are giant actin-binding proteins, orthologous to  
 RT Drosophila melanogaster muscle protein MSP-300.";  
 RL Genomics 80:473-481 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 6).  
 RX MEDLINE=22296983; PubMed=12408964;  
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12508121;  
 RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,  
 RA Cattolico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A.,  
 RA Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,  
 RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Crnaud C.,  
 RA Buelts T., Jallouin O., Friedlander L., Samson G., Brottier P.,  
 RA Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbasi N.,  
 RA Alich N., Boscut D., Dickhoff R., Dora M., Dubois I., Friedman C.,  
 RA Guayvenoux M., James R., Madan A., Mairay-Estrada B., Mangenot S.,  
 RA Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Traak B.,  
 RA Vacherie B., Bellemere C., Belser C., Beaudard-Comnet M.,  
 RA Bartol-Navel D., Bourard M., Bries-Silla S., Combette S.,  
 RA Dufosse-Laurent V., Ferron C., Lechaplais C., Louesse C., Muselet D.,  
 RA Magdelenat G., Pateau E., Petit E., Sirvain-Trukniewicz P., Trybou A.,  
 RA Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,  
 RA Dumont C., Guerin T., Haffray S., Hammadi R., Muanga J., Pellouin V.,  
 RA Robert D., Wundler E., Gauguier G., Roy A., Sainte-Marthe L.,  
 RA Verdier J., Verdier-Discala C., Hillier L.W., Fulton L., McPherson J.,  
 RA Matsuda F., Wilson R., Scarpelli C., Gypay G., Wincher P., Saurin W.,  
 RA Quetier F., Waterston R., Hood L., Weissbach J.;  
 RT "The DNA sequence and analysis of human chromosome 14.";  
 RL Nature 421:601-607 (2003).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM 8).  
 RX MEDLINE=22238857; PubMed=12477932;  
 RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan B., Moore K., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Ioshizuka S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6885  
 RP FROM N.A.  
 RX TISSUE=Splice, and Tongue;  
 RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.,  
 RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,  
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE OF 5709-6885 FROM N.A. (ISOFORM 2).  
 RX TISSUE=Brain;  
 RA MEDLINE=99246063; PubMed=10231032;  
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 6:63-70 (1999).  
 RN [9]  
 RP REVISIONS.  
 RX TISSUE=Brain;  
 RA MEDLINE=22158633; PubMed=12168954;  
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;  
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual  
 RT curation of 330 KIAA cDNA clones.";  
 RL DNA Res. 9:99-106 (2002).  
 RN [10]  
 RP SEQUENCE OF 5754-6885 FROM N.A.  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Poustka A.;  
 RT "Towards a catalog of human genes and proteins: sequencing and  
 RT analysis of 500 novel complete protein coding human cDNAs.";  
 RL Genome Res. 11:422-435 (2001).  
 CC -!- FUNCTION: Involved in the maintenance of nuclear organization and  
 CC structural integrity. Probable anchoring protein which tethers the  
 CC nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton  
 CC by interacting with the nuclear envelope and with F-actin in the  
 CC cytoplasm.  
 CC -!- SUBUNIT: Interacts with F-actin via its N-terminal domain.  
 CC -!- SUBCELLULAR LOCATION: Type IV membrane protein (potential). The  
 CC largest part of the protein is cytoplasmic, while its C-terminal  
 CC part is associated with the nuclear envelope, most probably the  
 CC outer nuclear membrane. Remains associated with the nuclear  
 CC envelope during its breakdown in mitotic cells.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=9;  
 CC Name=1;  
 CC IsoId=Q8WXH0-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8WXH0-2; Sequence=VSP\_007164, VSP\_007166;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:22:14 ; Search time 21.3962 Seconds  
(without alignments)  
3258.971 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_1\_221

Perfect score: 1125  
Sequence: 1 MAEGKAGAGLFAKQVOKK.....IGCVTFIFQNISNLRDVFYR 221

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25.\*  
1: sp archaeb.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organalle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1125	100.0	564	Q9UKN4	Q9ukn4 homo sapien
2	1114.5	99.1	565	Q86VW0	Q86vw0 homo sapien
3	1111.5	98.8	565	Q8UBW5	Q8ubw5 homo sapien
4	782.5	69.6	524	Q803I3	Q803i3 brachydanio
5	683	60.7	439	Q8ETH3	Q8eth3 homo sapien
6	683	60.7	490	Q8WH99	Q8wh99 homo sapien
7	662.5	58.9	478	Q7ZWP5	Q7zwp5 xenopus lae
8	625.5	55.6	250	Q8BXH2	Q8bxh2 mus musculu
9	625.5	55.6	686	Q7TGF7	Q7tgf7 mus musculu
10	620.5	55.2	385	Q8NFI3	Q8nfi3 homo sapien
11	620.5	55.2	695	Q8N4G0	Q8n4g0 homo sapien
12	449	35.9	414	Q8WQ54	Q8wq54 geodia cydo
13	429.5	38.2	602	Q8Y092	Q8y092 drosophila
14	405	36.0	461	Q21004	Q21004 caenorhabdi
15	228.5	20.3	273	Q8X0D7	Q8x0d7 neurospora
16	171.5	15.2	499	Q8NJ03	Q8nj03 neurospora

17	149	13.2	823	4	Q8IVY3	Q8ivy3 homo sapien
18	149	13.2	1315	4	Q9Y2L3	Q9y2l3 homo sapien
19	146	13.0	606	11	Q8ROY2	Q8roy2 mus musculu
20	144	12.8	253	11	Q9J108	Q9j108 mus musculu
21	141	12.5	137	4	Q9UQI2	Q9uqi2 homo sapien
22	129.5	11.5	1992	13	Q04834	Q04834 xenopus lae
23	127	11.3	217	2	Q9RSX5	Q9rsx5 streptococc
24	127	11.3	377	2	Q53475	Q53475 streptococc
25	120.5	10.7	639	4	Q96B40	Q96b40 homo sapien
26	120.5	10.7	652	4	Q9P2B4	Q9p2b4 homo sapien
27	120	10.7	400	16	Q8NZ79	Q8nz79 streptococc
28	119	10.6	1175	16	Q8XNW6	Q8xnw6 clostridium
29	118.5	10.5	650	4	Q8N9W4	Q8n9w4 homo sapien
30	118	10.5	399	17	Q58288	Q58288 pyrococcus
31	118	10.5	1186	16	Q895M7	Q895m7 clostridium
32	117.5	10.4	899	3	Q8NIZ0	Q8niz0 neurospora
33	116.5	10.4	1486	2	Q93IE8	Q93ie8 actinobacil
34	116	10.3	471	2	Q93SL9	Q93sl9 streptococc
35	116	10.3	1940	5	Q02456	Q02456 schistosoma
36	116	10.3	2771	5	Q26216	Q26216 plasmodium
37	115.5	10.3	2029	4	Q9C014	Q9c014 homo sapien
38	115.5	10.3	2069	4	Q9C013	Q9c013 homo sapien
39	115.5	10.3	2073	4	Q9C012	Q9c012 homo sapien
40	115	10.2	528	5	Q26589	Q26589 schistosoma
41	115	10.2	3641	13	Q7ZTH4	Q7zth4 fugu rubrip
42	114.5	10.2	532	2	Q55098	Q55098 streptococc
43	114.5	10.2	1320	11	Q9JK25	Q9jk25 rattus norv
44	114	10.1	1056	16	Q8REF7	Q8ref7 fusobacteri
45	114	10.1	1061	16	Q8DPM8	Q8dpm8 streptococc

ALIGNMENTS

RESULT 1  
Q9UKN4  
ID Q9UKN4 PRELIMINARY; PRT; 564 AA.  
AC Q9UKN4  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)  
DE Bridging integrator-2.  
GN BIN2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=20366138; PubMed=10903846;  
RX Ge K., Pendergast G.C.;  
RA "Bin2, a Functionally Nonredundant Member of the BAR Adaptor Gene  
RT Family.";  
RL Genomics 67:210-220 (2000).  
RR EMBL; AF14531; X54227.1; -  
DR GO; GO:0006897; P: endocytosis; IEA.  
DR GO; GO:0007268; P: synaptic transmission; IEA.  
DR InterPro; IPR003005; Amphiphysin.  
DR InterPro; IPR006632; BAR.  
DR InterPro; IPR004148; BAR\_dom.  
DR Pfam; PF03114; BAR; 1.  
DR PRINTS; PR01251; AMPHIPHYSIN.  
DR SMART; SM00721; BAR; 1.  
SQ SEQUENCE 564 AA; 61747 MW; 3CB3791A56CE53CB CRC64;

Query Match 100.0%; Score 1125; DB 4; Length 564;  
Best Local Similarity 100.0%; Pred. No. 5.8e-64;  
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAEGKAGAGLFAKQVOKKFSRAQSKVLQKLGKAVETKDERFEQASNFYQQAEHKL 60  
DB 1 MAEGKAGAGLFAKQVOKKFSRAQSKVLQKLGKAVETKDERFEQASNFYQQAEHKL 60



GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:22:14 ; Search time 54.6038 Seconds  
(without alignment)  
3258.971 Million cell updates/sec

Title: US-10-069-540A-2  
Perfect score: 2855  
Sequence: 1 MAEGKAGGAAGLPAKQVQKK.....NNLTAPFPQEVSTSENPL 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315519202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_ricet.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2851	99.9	564	Q9UKN4	Q9UKN4 homo sapien
2	2840.5	99.5	565	Q86VW0	Q86VW0 homo sapien
3	2837.5	99.4	565	Q2UBW5	Q2UBW5 homo sapien
4	1721	60.3	342	Q9NWK4	Q9NWK4 homo sapien
5	986	34.5	524	Q803I3	Q803I3 brachydanio
6	873	30.6	490	Q8WTH9	Q8WTH9 homo sapien
7	818.5	28.7	686	Q7TQF7	Q7TQF7 mus musculus
8	805.5	28.2	439	Q8BTH3	Q8BTH3 mus musculus
9	804.5	28.2	695	Q8N4G0	Q8N4G0 homo sapien
10	786.5	27.5	478	Q7ZWP5	Q7ZWP5 xenopus lae
11	743.5	26.0	385	Q8NFK3	Q8NFK3 homo sapien
12	705.5	24.7	250	Q8BXH2	Q8BXH2 mus musculus
13	555.5	19.5	602	Q9Y092	Q9Y092 drosophila
14	535.5	18.8	414	Q8WQ4	Q8WQ4 geodia cydo
15	463	16.2	461	Q21004	Q21004 caenorhabdi
16	256.5	9.0	137	Q9UQ12	Q9UQ12 homo sapien

ALIGNMENTS

RESULT 1

Q9UKN4  
ID Q9UKN4 PRELIMINARY; PRT; 564 AA.  
AC Q9UKN4  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 13, Last sequence update)  
DE Bridging integrator-2.  
GN BIN2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20366138; PubMed=10903846;  
RA Ge K., Prendergast G.C.;  
RT "Bin2, a Functionally Nonredundant Member of the BAR Adaptor Gene Family."  
RL Genomics 67:210-220(2000).  
DR EMBL; AF146531; AA54227.1;  
DR GO; GO:0006897; P: endocytosis; IEA.  
DR GO; GO:007268; P: synaptic transmission; IEA.  
DR InterPro; IPR003005; Amphiphysin.  
DR InterPro; IPR006632; BAR.  
DR Pfam; PF03114; BAR\_1.  
DR PRINTS; PR01251; AMPHIPHYSIN.  
DR SMART; SM00721; BAR; 1.  
SQ SEQUENCE 564 AA; 61747 MW; 3CB3791A56CE53CB CRC64;

Query Match 99.9%; Score 2851; DB 4; Length 564;  
Best Local Similarity 99.8%; Pred. No. 2.5e-139;  
Matches 563; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAEGKAGGAAGLPAKQVQKKFSAQKVLQKLGKAVETKDFEFQASNFYQQQAEHGKL 60  
DB 1 MAEGKAGGAAGLPAKQVQKKFSAQKVLQKLGKAVETKDFEFQASNFYQQQAEHGKL 60

```

QY 61 YKDLNFKLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWVNDLLWEDYEKLADQ 120
DB 61 YKDLNFKLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWVNDLLWEDYEKLADQ 120
QY 121 AVRTMEIYVAQSEIKERIAKGRKLVYDSARHHLAVQNAKDEAKTAKAEFFNKAQ 180
DB 121 AVRTMEIYVAQSEIKERIAKGRKLVYDSARHHLAVQNAKDEAKTAKAEFFNKAQ 180
QY 181 TVPEDLNQELLEELPILYNSRIGCVYTFQINISNLRDVYFREMKNHNLVEYMSKLEKQ 240
DB 181 TVPEDLNQELLEELPILYNSRIGCVYTFQINISNLRDVYFREMKNHNLVEYMSKLEKQ 240
QY 241 HSNKVFVVKGLSSRRSLVIPPVKTATVSSPLTSPSTLSKSESSEVSATEDLAP 300
DB 241 HSNKVFVVKGLSSRRSLVIPPVKTATVSSPLTSPSTLSKSESSEVSATEDLAP 300
QY 301 DAAQGEDNSEIKELLEEEIEKEGSEASSEDEPLPACNGPAQAPSPPTTERAKSQEEV 360
DB 301 DAAQGEDNSEIKELLEEEIEKEGSEASSEDEPLPACNGPAQAPSPPTTERAKSQEEV 360
QY 361 LPSSTTPSPGGALSPSGQSPSSATEVLRTRTASEGSEQPKKRASIORTSAPPSPPTT 420
DB 361 LPSSTTPSPGGALSPSGQSPSSATEVLRTRTASEGSEQPKKRASIORTSAPPSPPTT 420
QY 421 ATASPRPSSGNTPSPSTASGGSPSPRASLTGTGTASPTSLVSPNPEPEKPVRTPEA 480
DB 421 ATASPRPSSGNTPSPSTASGGSPSPRASLTGTGTASPTSLVSPNPEPEKPVRTPEA 480
QY 481 KENENIHQNPEELCTSPMTSQVASEPGEAKMEDKKNKLISADSGDQLOVSM 540
DB 481 KENENIHQNPEELCTSPMTSQVASEPGEAKMEDKKNKLISADSGDQLOVSM 540
QY 541 VPENNNLTAPEQEEVSTSENQOL 564
DB 541 VPENNNLTAPEQEEVSTSENQOL 564

RESULT 2
Q86VVO ID Q86VVO PRELIMINARY; PRT; 565 AA.
AC Q86VVO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Breast cancer associated protein BRAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feilsgold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussdin T.B., Ioshizuka S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek A., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

```

```

[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047686; AAH47686.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PRO1251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 565 AA; 61875 MW; BIDIFB89A7707EA3 CRC64;

Query Match 99.5%; Score 2840.5; DB 4; Length 565;
Best Local Similarity 99.6%; Pred. No. 8.9e-139;
Matches 563; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAEGKAGGAAGLFAKQVKKFSRAQEKVLQKLGKAVETKDERFEQASNFYQQAEGHKL 60
DB 1 MAEGKAGGAAGLFAKQVKKFSRAQEKVLQKLGKAVETKDERFEQASNFYQQAEGHKL 60
QY 61 YKDLNFKLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWVNDLLWEDYEKLADQ 120
DB 61 YKDLNFKLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWVNDLLWEDYEKLADQ 120
QY 121 AVRTMEIYVAQSEIKERIAKGRKLVYDSARHHLAVQNA-KKDEAKTAKAEFFNKA 179
DB 121 AVRTMEIYVAQSEIKERIAKGRKLVYDSARHHLAVQNA-KKDEAKTAKAEFFNKA 180
QY 180 QTVFEDLNQELLEELPILYNSRIGCVYTFQINISNLRDVYFREMKNHNLVEYMSKLEK 239
DB 181 QTVFEDLNQELLEELPILYNSRIGCVYTFQINISNLRDVYFREMKNHNLVEYMSKLEK 240
QY 240 QHSNKFVVKGLSSRRSLVIPPVKTATVSSPLTSPSTLSKSESSEVSATEDLA 299
DB 241 QHSNKFVVKGLSSRRSLVIPPVKTATVSSPLTSPSTLSKSESSEVSATEDLA 300
QY 300 FDDAQGEDNSEIKELLEEEIEKEGSEASSEDEPLPACNGPAQAPSPPTTERAKSQEE 359
DB 301 FDDAQGEDNSEIKELLEEEIEKEGSEASSEDEPLPACNGPAQAPSPPTTERAKSQEE 360
QY 360 VLPSTTPSPGGALSPSGQSPSSATEVLRTRTASEGSEQPKKRASIORTSAPPSPPTT 419
DB 361 VLPSTTPSPGGALSPSGQSPSSATEVLRTRTASEGSEQPKKRASIORTSAPPSPPTT 420
QY 420 RATASPRPSSGNTPSPSTASGGSPSPRASLTGTGTASPTSLVSPNPEPEKPVRTPE 479
DB 421 RATASPRPSSGNTPSPSTASGGSPSPRASLTGTGTASPTSLVSPNPEPEKPVRTPE 480
QY 480 AKENENIHQNPEELCTSPMTSQVASEPGEAKMEDKKNKLISADSGDQLOVSM 539
DB 481 AKENENIHQNPEELCTSPMTSQVASEPGEAKMEDKKNKLISADSGDQLOVSM 540
QY 540 MVPENNNLTAPEQEEVSTSENQOL 564
DB 541 MVPENNNLTAPEQEEVSTSENQOL 565

RESULT 3
Q9UBW5 ID Q9UBW5 PRELIMINARY; PRT; 565 AA.
AC Q9UBW5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Breast cancer associated protein BRAP1.
GN BRAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Miki Y., Saito H.;
RT "Genomic structure and chromosome location of the BRP1 gene.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032710; BAA88125.1; JOINED.
DR EMBL; AB032699; BAA88125.1; JOINED.
DR EMBL; AB032700; BAA88125.1; JOINED.
DR EMBL; AB032701; BAA88125.1; JOINED.
DR EMBL; AB032702; BAA88125.1; JOINED.
DR EMBL; AB032703; BAA88125.1; JOINED.
DR EMBL; AB032704; BAA88125.1; JOINED.
DR EMBL; AB032705; BAA88125.1; JOINED.
DR EMBL; AB032706; BAA88125.1; JOINED.
DR EMBL; AB032707; BAA88125.1; JOINED.
DR EMBL; AB032708; BAA88125.1; JOINED.
DR EMBL; AB032709; BAA88125.1; JOINED.
DR EMBL; AB032698; BAA88108.1; -.
DR Genew; HGNC:1053; BIN2.
DR GO; GO:0006897; P:endocytosis; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 565 AA; 61902 MW; 28D1FB89A7779860 CRC64;

Query Match 99.4%; Score 2837.5; DB 4; Length 565;
Best Local Similarity 99.5%; Pred. No. 1.3e-138;
Matches 562; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAEGKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDRFRFQSASNFYQQAEGHKL 60
DB 1 MAEGKAGGAGLFAKQVQKFSRAQEKVLQKLGKAVETKDRFRFQSASNFYQQAEGHKL 60
QY 61 KYDLKNFLSAVKMHESSKRVSETLQIYSSEWDGHEELKAIWNNDLLMEDYEELADQ 120
DB 61 KYDLKNFLSAVKMHESSKRVSETLQIYSSEWDGHEELKAIWNNDLLMEDYEELADQ 120
QY 121 AVRTMEIYVAQSEIKERIAKGRKLVVDYDSARHLEAVQNA-KDDEAKTAKAEFEENKA 179
DB 121 AVRTMEIYVAQSEIKERIAKGRKLVVDYDSARHLEAVQNA-KDDEAKTAKAEFEENKA 180
QY 180 QTVFEDLNQELLEELPILYNSRIGCYVTIFQNIENLRDVFYREMSKLNHLYEVMKLEK 239
DB 181 QTVFEDLNQELLEELPILYNSRIGCYVTIFQNIENLRDVFYREMSKLNHLYEVMKLEK 240
QY 240 QHSNKVFVVKGLSSSRSLVSPVTRATVSSPLTSPSTSLKSESESVSATEDLA 299
DB 241 QHSNKVFVVKGLSSSRSLVSPVTRATVSSPLTSPSTSLKSESESVSATEDLA 300
QY 300 PDAAQGEDNSETKELLEEEIEKEGSEASSEEDDPLPACNPAQAQPSPTTERAKSQEE 359
DB 301 PDAAQGEDNSETKELLEEEIEKEGSEASSEEDDPLPACNPAQAQPSPTTERAKSQEE 360
QY 360 VLPSTTTPSGGALSPGQOPSSATEVVLRTTRTASGSEQPKKASTQRTASPPSPPPP 419
DB 361 VLPSTTTPSGGALSPGQOPSSATEVVLRTTRTASGSEQPKKASTQRTASPPSPPPP 420
QY 420 RATASPRSSGNIIPSSPTASGGGSPSPRASLGTTGTASPTSLSEVSNPPEPKPVRTPE 479
DB 421 RATASPRSSGNIIPSSPTASGGGSPSPRASLGTTGTASPTSLSEVSNPPEPKPVRTPE 480
QY 480 AKENENIHNQPEELCTSPTLMTSQVASEPGEAKMEDKONKLI SADSEGDQLQVS 539
DB 481 AKENENIHNQPEELCTSPTLMTSQVASEPGEAKMEDKONKLI SADSEGDQLQVS 540
QY 540 MYPENNNTAPEPQEVSTSENPOL 564
DB 541 MYPENNNTAPEPQEVSTSENPOL 565

RESULT 4
Q9NWK4 PRELIMINARY; PRT; 342 AA.
ID Q9NWK4;
AC Q9NWK4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ20776.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000783; BAA31376.1; -.
KW Hypothetical protein.
SQ SEQUENCE 342 AA; 35986 MW; 8DD6B908D9A28DBD CRC64;

Query Match 60.3%; Score 1721; DB 4; Length 342;
Best Local Similarity 99.7%; Pred. No. 2.6e-81;
Matches 341; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 223 MSKLNHNLYEVMKLEKHQHNKVVVKGSLSSRRSLVSPVTRATVSSPLTSPST 282
DB 1 MSKLNHNLYEVMKLEKHQHNKVVVKGSLSSRRSLVSPVTRATVSSPLTSPST 60
QY 283 LSLKSESVSATDLDAPDAQAQEDNSEIKELLEEEIEKEGSEASSEEDDPLPACNGP 342
DB 61 LSLKSESVSATDLDAPDAQAQEDNSEIKELLEEEIEKEGSEASSEEDDPLPACNGP 120
QY 343 AAQPSPTTERAKSQEVLFPSTTTPSGGALSPGQOPSSATEVVLRTTRTASGSEQPKK 402
DB 121 AAQPSPTTERAKSQEVLFPSTTTPSGGALSPGQOPSSATEVVLRTTRTASGSEQPKK 180
QY 403 RASTQRTASPPSPPPPRTASPPSSGNTSPSSGTPSSGSPSPRASLGTTGTASPTSL 462
DB 181 RASTQRTASPPSPPPPRTASPPSSGNTSPSSGTPSSGSPSPRASLGTTGTASPTSL 240
QY 463 EVSPNPPEPKPVRTPEAKENIHNQPEELCTSPTLMTSQVASEPGEAKMEDKON 522
DB 241 EVSPNPPEPKPVRTPEAKENIHNQPEELCTSPTLMTSQVASEPGEAKMEDKON 300
QY 523 KLISADSEGDQLQVSNVPENNNLTAPEPQEVSTSENPOL 564
DB 301 KLISADSEGDQLQVSNVPENNNLTAPEPQEVSTSENPOL 342

RESULT 5
Q803I3 PRELIMINARY; PRT; 524 AA.
ID Q803I3;
AC Q803I3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to bridging integrator 2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044473; AAH44473.1; -.

```

[illegible]



Db 247 GAPSDSGPLRIAKTTPPEEPFSPCTPTASPTSENHTLAPASPAPARPSPSQT----- 299  
 Qy 301 DAAQGDNDSEIKELLEEEIEKGGSEASSSEDPLP--ACNGPAQAOPRPTTERAKSOE 358  
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
 Db 298 --RKGPVPVPLKVTPTKELQE--NIISFPEDNFVEI SVITPSQNE----VPEYKKEE 349  
 Qy 359 EVL-----PSSTTPSPGALSPSQPPS-----SATEVVLRTTRTASEGS-----EQ 400  
 Db 350 TLLDLDFDPFKPEVTVPAGSAGVTHSPMSOTLPMDLTWTSTDLV---QPASGGSGFNFTQP 406  
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
 Qy 401 KKASI-----ORTSAPSRPPRPPRATASPRESSGNIPSSPTASGGGSPTSPT 447  
 Db 407 QDTSLFTMGTDSMICNLAESEQAPPTPKAEBPLAATPVG----- 449  
 Qy 448 RASLGTTGATSPRTSLSEVSNPEPPKPVRTPEAKENINIHONPEELCTSPMTMTSQVAS 507  
 Db 450 -LDLGMDTRA-----EEPVEEA VII PGADADA AVGT-----LVSAABG 486  
 Qy 508 EGEKAKKVEDK-----EKDNKL I SADSSGODQL OYSWMPENNNTLAPBPQEVST 559  
 Db 487 APCEEAETATVPAGEGVGLEAKIGITGETTEGAESAQ-----PEABEALTVPQEKVIP 542

RESULT 10  
 Q7ZWPS PRELIMINARY; PRT; 478 AA.  
 AC Q7ZWP5;  
 RC Q7ZWP5;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Similar to bridging integrator 1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Eureleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 CC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; BC046852; AAA46852.1; -.  
 DR GO; GO:0006897; P:endocytosis; IEA.  
 DR CO; GO:0007269; P:synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR PRINTS; PRO0452; SH3DOMAIN.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR SEQUENCE 478 AA; 53445 MW; 042A1A94D59095A8 CRC64;

Query Match 27.5%; Score 786.5; DB 13; Length 478;  
 Best Local Similarity 42.6%; Pred No. 5; e=-33;  
 Matches 190; Conservative 59; Mismatches 128; Indels 59; Gaps 11;

Qy	1	M AEGKAGGAAGLF AKVQVKXFSRAQKV LQKL GKAVETKDERFEOSANFYQQQAEHGHL	60
Db	1	N AEELKGVSAGCLASNMQKLL TRAQKV LQKL GKADETDEQ FESCQNFKQLTEGSKL	60
Qy	61	Y KDLKNFLSA VKVGHSS KEVSTLQEI YSSEWDGHEELKATVMNDLLWEDYEELKADQ	120
Db	61	Q XDLERAYLATV KMAHEASKKL TELL DVTEPDMPGREENTNKTAENN DILLWTDYHQKLV DQ	120
Qy	121	A VRTWEITYVAQFS IKERIAKRGRKLVDYDSARHHLEAVQNA-KKDEAKTKAEDEFNKA	179
Db	121	A LLTWDTYLGGFPDI KSIAKRGRKLVDPSARHHFPESLNQAKKOETKIKAAREDDLKS	180

```

QY 180 QTVFEDLNOELLELPILYNSRIGCVYTFONISNLRDVFYREMSKLNHLYEVMKLEK 239
Db 181 QKVFEMVNDLQEEPLSLWNSRVGYVTFQSIAGLENFHFKMGKHLQMLNVMWVLES 240
QY 240 Q-----HNNKVVYVYKGL--SSSSRRSLV-----GPPVRTATVGSPLT 275
Db 241 EGNLSTNSSSVKQVPSLKHNTNSNRKKNWLFSLRIKKGSHCHISESTPTKNTNPSPPA 300
QY 276 SPTSSTLSLKSESSESVATEDL-APDAAGEDNSEIKELLEEEIEKEGSEASSSEEDD 334
Db 301 SGPFAV-----SETKVHDFEAPSA-----VEETSVSAISPSPOFVE 341
QY 335 PLPACNGPAQAQSPPTTERAKSQEELPSTTPSPGALSPGQSSSATEWLRTAS 394
Db 342 PAPA-----PVSVNEPAKETEVPESAPPPEVDPKSPSSLPA-----VVVEIFSA 388
QY 395 EGSQPKKRAIQTSAAPSPRPPPP 419
Db 389 VNGTVE-----SSPPKPEMP 404

RESULT 11
Q8NFL3
ID Q8NFL3 PRELIMINARY; PRT; 385 AA.
AC Q8NFL3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Amphiphysin I variant Nt2 (Fragment).
GN AMPH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Terada Y., Teutsui K., Sano K., Hosoya O., Ohtsuki H., Tokunaga A.,
RA Tsutsui K.;
RT "Novel splice variants of amphiphysin I are expressed in retina.";
RL FEBS Lett. 519:185-190(2002).
DR EMBL; AF498097; AAM44811.1; -.
DR GO; GO:0006897; P:endocytosis; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR003017; Amphiphysin_1.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR ProDom; PD003208; Amphiphysin_1; 1.
DR SMART; SM00721; BAR; 1.
FT NON_TER 385
SQ SEQUENCE 385 AA; 43753 MW; 499AB0A3136711AA CRC64;

Query Match 26.0%; Score 743.5; DB 4; Length 385;
Best Local Similarity 41.9%; Pred. No. 7.4e-31;
Matches 164; Conservative 77; Mismatches 113; Indels 37; Gaps 9;

QY 11 GLPAQVQKKFSRAQEKVLQKLGKAVETKDERPEQASNFYQQAEGHKLKXDLKNFLSA 70
Db 7 GIFAKNVQKLNRAQEKVLQKLGKADETKDEQFEYVQNFKQEAEGTLQRELGYLAA 66
QY 71 VVMHESKRVSETLQETYSSEWDGHEELKAIWNNDLLWEDYEKLADQAVRTMEIYA 130
Db 67 IKGQEAEMKLTESLHEVPEPDWYGRDVGMVGEKCDVLWEDFHQKLVDSLLTDTYLG 126
QY 131 QFSEIKERIAKRGKLVNDYDSARHLEAVQNAK-KDEAKTAKAEFEFNKAQTVFEDLNOE 189
Db 127 QFPDIKNRIAKRSKLVNDYDSARHLEALQSSKEKDSRISKAEEFQKAQKVFEEFNVD 186
QY 190 LLEELPILYNSRIGCVYTFONISNLRDVFYREMSKLNHLYEVMKLEKQHSNKVYVVK 249
Db 187 LQEEPLSLWNSRVGYVTFONISNLRDVFYREMSKLNHLYEVMKLEKQHSNKVYVVK 246
QY 250 G 250
Db 247 G 247

```

```

Db 187 LQEEPLSLWNSRVGYVTFONISNLRDVFYREMSKLNHLYEVMKLEKQHSNKVYVVK 246
QY 250 GLSSSSRRSLVISPPVTATVSSPLTSPSTLSKSESSESVATEDLAP----- 300
Db 247 G-----APRMCRDSCGLRIAKTPSP--PBPSPSL-----PSPTASPNHTLAPASPAPAPR 295
QY 301 --DAAQGEDNSEIKELLEEEIEKEGSEASSSEEDDPLP--ACNGPAQAQSPPTTERAK 355
Db 296 SPSQTRKGPVPPPLPKVPTPKELQOE--NIISFEENFVPEISVITTPSQNE-----VPEVK 349
QY 356 SQEEVL-----PSSTTPSPGALSPSQSPSS 381
Db 350 KEETILLDLDFFPKPEVTPAGSAGVTHSPMS 380

RESULT 12
Q8BXH2
ID Q8BXH2 PRELIMINARY; PRT; 250 AA.
AC Q8BXH2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to MYC box dependent interacting protein 1.
GN AMPH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK047144; BAC32971.1; -.
DR MGD; MGI:103574; Amph.
DR GO; GO:0006897; P:endocytosis; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR Pfam; PF03114; BAR; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR SMART; SM00721; BAR; 1.
SQ SEQUENCE 250 AA; 29125 MW; 568A3D55CC6C37CE CRC64;

Query Match 24.7%; Score 705.5; DB 11; Length 250;
Best Local Similarity 54.8%; Pred. No. 4e-29;
Matches 132; Conservative 56; Mismatches 52; Indels 1; Gaps 1;

QY 11 GLPAQVQKKFSRAQEKVLQKLGKAVETKDERPEQASNFYQQAEGHKLKXDLKNFLSA 70
Db 7 GIFAKNVQKLNRAQEKVLQKLGKADETKDEQFEYVQNFKQEAEGTLQRELGYLAA 66
QY 71 VVMHESKRVSETLQETYSSEWDGHEELKAIWNNDLLWEDYEKLADQAVRTMEIYA 130
Db 67 IKGQEAEMKLTESLHEVPEPDWYGRDVGMVGEKCDVLWEDFHQKLVDSLLTDTYLG 126
QY 131 QFSEIKERIAKRGKLVNDYDSARHLEAVQNAK-KDEAKTAKAEFEFNKAQTVFEDLNOE 189
Db 127 QFPDIKNRIAKRSKLVNDYDSARHLEALQSSKEKDSRISKAEEFQKAQKVFEEFNVD 186
QY 190 LLEELPILYNSRIGCVYTFONISNLRDVFYREMSKLNHLYEVMKLEKQHSNKVYVVK 249
Db 187 LQEEPLSLWNSRVGYVTFONISNLRDVFYREMSKLNHLYEVMKLEKQHSNKVYVVK 246
QY 250 G 250
Db 247 G 247

```

RESULT 13  
 QY092 PRELIMINARY; PRT; 602 AA.  
 AC QY092; 12, Created)  
 DT 01-NOV-1999 (TRENDELrel. 12, Last sequence update)  
 DT 01-JUN-1999 (TRENDELrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TRENDELrel. 24, Last annotation update)  
 DE Amphiphysin protein (LID19810P).  
 GN AMPH OR CGS604.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter Z.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Fianknoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dalke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Deicher A., Beng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Flocker A., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Talali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kimbel B.E., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20076404; PubMed=10607911;  
 RA Razaq A., Su Y., Mehren J.E., Mizuguchi K., Jackson A.P., Gay N.J.,  
 RA O'Kane C.J.,  
 RT "Characterization of the gene for Drosophila amphiphysin."  
 RL Gene 241:167-174 (2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Y.Q., Xue H., Broadie K.S.,  
 RT "Dissecting the functions of amphiphysin."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nurco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.,  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMEL; AF171225; AAF58486.1; -;  
 DR EMEL; AF171225; AAD46684.1; -;  
 DR EMEL; AJ242855; CAB45188.1; -;  
 DR EMEL; AY061278; AAL28826.1; -;  
 DR HSPF; P27986; 1EHT.  
 DR Flybase; FBgn0027356; Amph.  
 DR GO; GO:0005886; C:plasma membrane; IDA.  
 DR GO; GO:0045202; C:synaptic junction; IDA.  
 DR GO; GO:0007269; P:neurotransmitter secretion; NAS.  
 DR GO; GO:0008104; P:protein localization; IMP.  
 DR GO; GO:0006937; P:regulation of muscle contraction; IMP.  
 DR GO; GO:0045313; P:rhabdosome membrane biogenesis; IMP.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR dom.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF03114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR SH3 domain. 602 AA; 65901 MW; 86D1D4CEAE1B8F31 CRC64;  
 SQ SEQUENCE

Query Match 19.5%; Score 555.5; DB 5; Length 602;  
 Best Local Similarity 28.4%; Pred. No. 6.3e-21;  
 Matches 169; Conservative 90; Mismatches 240; Indels 97; Gaps 15;

QY 12 LFAQVQKSRQAEKVLQKLGAVETQERPEQASNFVQQAEGHKLKDLKXLSAV 71  
 DB 8 MLAKSVQKHAGRAKELQNLGKVDRTADEIFDHLNFFNQQAASNRUQKGFNNYIRCV 67  
 QY 72 KVMHSSKRVSETLQETISYSEWDGHEKLAIVWNLLWEDYEKLADQAVTMEIYVAQ 131  
 DB 68 RAAQAASKLWDSVCEIYEPQWSSYDALQATQAGSESLWADPAHKLGDQVLPNTYTQ 127  
 QY 132 FSEIKERIAGKRLVDYDSARH---HLRAVQAKQDEAKTAKAEEFNKAQVTFEDLNQ 188  
 DB 128 FPEKKXVEKENRKLIDYDQGRHSFQNLQANAKRDXDKLTGREGLEAEARTYEILNT 187  
 QY 189 ELLEELPLIYNSRIGCVVTFQNISNLRDVFREMSKLNHNLVEVMVKL--EKQHSNKVF 246  
 DB 188 ELHDELPAIDSRLLFLVTLQTLFATEQVFNHETAKIYSELEAIVDKLATESQSGSNT- 246  
 QY 247 VVKGLSSSSRRSLVISPPVR-----TATVSSPLTSPSTPSLTSLKSESE 290  
 DB 247 LRKQTSNPITKSPVQSPVFNKLNANINSNYQITNGSGSLANSPTSTSSLOEPRFD 306  
 QY 291 SVSATEDLAPDA-----AQGEDNSEIKELLEEEIEKEGSEASSSEDDPLPACNGPAQA 345  
 DB 307 SVSPTPPTSPESPAALVSATPSGPV-----ENGVTTKSLERFELSGLNASAKA 355  
 QY 346 ----QPSFTTERAKSQBEVLPSSTTPPGGALSPGQSSSSATEVLRTRT-----ASEG 396  
 DB 356 TTTTQTSPTEDKAVVSAVKPSET---EGAAVAASVTPAPPATPAQINGNNNEPSIVKEG 412  
 QY 397 SEQPKKASQRTSAPPSPRPPTASPPSCGNIPSSPTAGGSGSPSPRAS-----450  
 DB 413 GKQPKEL-----PSTTSNAEAAEAAANNGNSIEEHKQKGLGNDTIVTATETVTVT 463  
 QY 451 ----LGTGT-----ASPTSLSEVSPNPEPPEKVPRTPEAKENENIENQNPPE-----493  
 DB 464 QHSVTSTDTNIVTISTNTDTTKTSTGTSQKRPVFPVVRNHSVNNL--NKKPFEDDDDR 522





```

Db 270 SDNSSLRQ-----TPSPS-----VLPS-- 286
QY 366 TPSPGALSPSQPSSSATEVVLRTETATSESGSEQPKKEASIQRTSAPSRPPPPRATASP 425
Db 287 -FVP-----APMAKPREVVSMANSTNPPDDDDAEHDKAT-----PTDVEEKFEAKVYP 335
QY 426 RPSSGNIPSPSTASGGSPSPRASLGCTATSP-----RTSLEVSP---NPEPPEKPV 475
Db 336 KLNAAQTAAAEQAIAAAKKEKES-----NPPDDEDESTEVEEAPKKKDPPTPKPL 389
QY 476 RTEAKENENIHQNPEEL--CTS-----PTLMTSQVASEPGEAKKQEDKXDNKL 524
Db 390 -----DGITNEKRTLYFVTSTHNYKAVDTDELSFEQGVQIKVIEANEDDQL 436

```

Search completed: March 4, 2004, 17:29:50  
 Job time : 58.6038 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:19:24 ; Search time 72.5656 Seconds  
(without alignments)  
2196.038 Million cell updates/sec

Title: US-10-069-540A-2  
Perfect score: 2855  
Sequence: 1 MAEGKAGAGLFAKQVKK.....NLTAPPEQEVSTENPQL 564

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2855	100.0	564	AAB62100	Aab62100 Human bri
2	2794.5	97.9	594	ABG20887	Abg20887 Novel hum
3	2794.5	97.9	594	ABG20324	Abg20324 Novel hum
4	2787.5	97.6	594	ABG22457	Abg22457 Novel hum
5	2765.5	96.9	572	ABG21411	Abg21411 Novel hum
6	1988.5	63.6	414	ABG13005	Abg13005 Novel hum
7	1846	64.7	425	ABG22456	Abg22456 Novel hum
8	1788	62.6	960	ABG20886	Abg20886 Novel hum
9	1788	62.6	960	ABG20323	Abg20323 Novel hum
10	858	30.1	588	ADD45951	Add45951 Rat Prote
11	858	30.1	588	ADE60848	Ade60848 Rat Prote
12	858	30.1	588	ADE60840	Ade60840 Rat Prote
13	858	30.1	588	ADE60844	Ade60844 Rat Prote
14	858	30.1	588	ADE60852	Ade60852 Rat Prote
15	851.5	29.8	593	ADE60842	Ade60842 Human Pro
16	851.5	29.8	593	ADD45953	Add45953 Human Pro
17	851.5	29.8	593	ADE60850	Ade60850 Human Pro
18	851.5	29.8	593	ADE60846	Ade60846 Human Pro
19	851.5	29.8	593	ADE60854	Ade60854 Human Pro
20	829	28.0	482	ABU89771	Abu89771 Protein d
21	815.5	28.6	434	AAW05391	Aaw05391 Mouse SH3
22	815.5	28.6	683	ADD44887	Add44887 Rat Prote
23	804.5	28.2	695	AAU75110	Aau75110 Synaptic
24	804.5	28.2	695	ADD44889	Add44889 Human Pro
25	803.5	28.1	695	AAE22091	Aae22091 Human amp

26	793.5	27.8	451	2	AAW47295	Aaw47295 Human Bin
27	793.5	27.8	451	2	AAW94504	Aaw94504 Human Bin
28	792	27.7	451	2	AAW06602	Aaw06602 Human Box
29	792	27.7	451	2	AAW36881	Aaw36881 Human Bin
30	767	26.9	252	4	ABG22101	Abg22101 Bar domai
31	763.5	26.7	678	4	ABG22453	Abg22453 Novel hum
32	763.5	26.7	721	4	ABG20321	Abg20321 Novel hum
33	763.5	26.7	721	4	ABG20884	Abg20884 Novel hum
34	720.5	25.2	453	6	ABR69631	AbR69631 Human CGD
35	712	24.9	233	4	ABG22454	Abg22454 Novel hum
36	675	23.6	404	2	AAW05392	Aaw05392 Human SH3
37	625.5	21.9	130	4	ABG22449	Abg22449 Novel hum
38	625.5	21.9	130	4	ABG20881	Abg20881 Novel hum
39	625.5	21.9	130	4	ABG21407	Abg21407 Novel hum
40	625.5	21.9	130	4	ABG20318	Abg20318 Novel hum
41	555.5	19.5	602	4	ABG63948	Abg63948 Drosophil
42	546	19.1	106	3	ABG33907	Abg33907 Human sec
43	527	18.5	108	5	ABB06161	Abb06161 Human NS
44	527	18.5	137	4	ABG21410	Abg21410 Novel hum
45	367.5	12.9	82	4	ABG20317	Abg20317 Novel hum

## ALIGNMENTS

RESULT 1  
AAB62100  
ID AAB62100 standard; protein; 564 AA.

XX AC AAB62100;  
XX DT 29-MAY-2001 (first entry)  
XX DE Human bridging integrator-2 (Bin2) protein.  
XX KW Bridging integrator-2 protein; Bin2; cancer; leukemia; blood disorder;  
XX KW hyperplastic disease; cytostatic; cell growth regulator;  
XX KW chromosome 4q22.1.  
XX OS Homo sapiens.  
XX FH Key  
XX FT Peptide  
XX FT Peptide  
XX FT Peptide  
XX PN WO200116158-A2.  
XX PD 08-MAR-2001.  
XX PF 30-AUG-2000; 2000WO-US023723.  
XX PR 31-AUG-1999; 99US-0151554P.  
XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
XX PI Prendergast GC, Ge K;  
XX DR WPI; 2001-235087/24.  
XX DR N-ESDB; AAF57268.  
XX PT New bridging integrator-2 (Bin2) protein and nucleic acid molecules,  
XX PT useful for regulating cell growth, and for diagnosing or treating  
XX PT conditions associated with inappropriate expression of Bin2, e.g. cancers  
XX PT or hepatocarcinoma.  
XX PS Claim 1; Fig 1A-C; 62pp; English.  
XX CC This represents a human bridging integrator-2 (Bin2) protein. The Bin2  
XX CC protein can be expressed by standard recombinant methodology. The Bin2  
XX CC proteins or peptides are useful in regulating cell growth, cell survival,  
XX CC differentiation, endocytosis and actin organization. These peptides or

CC proteins are also useful for diagnosing or treating conditions associated  
 CC with inappropriate expression of Bin2, e.g. liver cancer, breast cancer,  
 CC hepatocarcinoma, myeloid and lymphoid leukemias or blood disorders. The  
 CC peptides, nucleic acid sequences or anti-Bin2 antibodies are useful for  
 CC diagnosing inappropriate expression of Bin2. Bin2 is also useful for  
 CC treating disorders associated with excessive Bin1 levels, e.g. liver,  
 CC colorectal, prostate or breast cancers, epithelia cell cancers, melanoma,  
 CC or hyperplastic disease states.

XX  
 SQ Sequence 564 AA;

Query Match 100.0%; Score 2855; DB 4; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-167;  
 Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEGKAGGAGLPAKQVKKFSAQKVKLQKGAETKDERFQSSANFYQQQAEHKL 60  
 DB 1 MAEGKAGGAGLPAKQVKKFSAQKVKLQKGAETKDERFQSSANFYQQQAEHKL 60

QY 61 YKDLKNFLSAVKVMHSSKRVSETLQETYSSEWDGHEELKAIWVNDLLWEDYEELADQ 120  
 DB 61 YKDLKNFLSAVKVMHSSKRVSETLQETYSSEWDGHEELKAIWVNDLLWEDYEELADQ 120

QY 121 AVRTMEIYVAQSEIKERIAKRGKLVYDSARHLEAVQNAKDEAKTAKAEFEFNKA 180  
 DB 121 AVRTMEIYVAQSEIKERIAKRGKLVYDSARHLEAVQNAKDEAKTAKAEFEFNKA 180

QY 181 TVFEDLNQELLELPILYNSRIGCVYTFQNTISNLRDVFYREMSKLNHLYEVMSKLEKQ 240  
 DB 181 TVFEDLNQELLELPILYNSRIGCVYTFQNTISNLRDVFYREMSKLNHLYEVMSKLEKQ 240

QY 241 HSNKVFVVKGLSSSRSLVISPVTATVSPSTLSKSESVSATEDLAP 300  
 DB 241 HSNKVFVVKGLSSSRSLVISPVTATVSPSTLSKSESVSATEDLAP 300

QY 301 DAAQGEDNSEIKELLEEEIEKESGSEASSEDDEPLPACNGPAQAQSPPTTERAKSQEEV 360  
 DB 301 DAAQGEDNSEIKELLEEEIEKESGSEASSEDDEPLPACNGPAQAQSPPTTERAKSQEEV 360

QY 361 LPSSTTPSPGGLSPSGQSSSATEVLRTRTASGSGQPKKRASIQRTASPPRPPPPR 420  
 DB 361 LPSSTTPSPGGLSPSGQSSSATEVLRTRTASGSGQPKKRASIQRTASPPRPPPPR 420

QY 421 ATASPRPSSGNTPSSPTASGGSPSPRASLTGTASPTSLVSPNPEPEKPVRTPEA 480  
 DB 421 ATASPRPSSGNTPSSPTASGGSPSPRASLTGTASPTSLVSPNPEPEKPVRTPEA 480

QY 481 KENENIHQNPELCTSPMTSQVASEPGEAKMEDKDKNKLISADSSGQDLQVSM 540  
 DB 481 KENENIHQNPELCTSPMTSQVASEPGEAKMEDKDKNKLISADSSGQDLQVSM 540

QY 541 VPENNNLTAPEPQEEVSTSENPL 564  
 DB 541 VPENNNLTAPEPQEEVSTSENPL 564

RESULT 2  
 ID ABG20887  
 XX ABG20887 standard; protein; 594 AA.  
 AC ABG20887;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #20878.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX

PD 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Dmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS85074.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51246; 103pp; English.  
 CC  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have application in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 594 AA;

Query Match 97.9%; Score 2794.5; DB 4; Length 594;  
 Best Local Similarity 98.1%; Pred. No. 5e-163;  
 Matches 554; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MAEGKAGGAGLPAKQVKKFSAQKVKLQKGAETKDERFQSSANFYQQQAEHKL 60  
 DB 30 MAEGKAGGAGLPAKQVKKFSAQKVKLQKGAETKDERFQSSANFYQQQAEHKL 89

QY 61 YKDLKNFLSAVKVMHSSKRVSETLQETYSSEWDGHEELKAIWVNDLLWEDYEELADQ 120  
 DB 90 YKDLKNFLSAVKVMHSSKRVSETLQETYSSEWDGHEELKAIWVNDLLWEDYEELADQ 149

QY 121 AVRTMEIYVAQSEIKERIAKRGKLVYDSARHLEAVQNA-KDEAKTAKAEFEFNKA 179  
 DB 150 AVRTMEIYVAQSEIKERIAKRGKLVYDSARHLEAVQNAKDEAKTAKAEFEFNKA 209

QY 180 QTVFEDLNQELLELPILYNSRIGCVYTFQNTISNLRDVFYREMSKLNHLYEVMSKLEK 239  
 DB 210 QTVFEDLNQELLELPILYNSRIGCVYTFQNTISNLRDVFYREMSKLNHLYEVMSKLEK 269

QY 240 QHSNKVFVVKGLSSSRSLVISPVTATVSPSTLSKSESVSATEDLA 299  
 DB 270 QHSNKVFVVKGLSSSRSLVISPVTATVSPSTLSKSESVSATEDLA 329

QY 300 PDAQGEDNSEIKELLEEEIEKESGSEASSEDDEPLPACNGPAQAQSPPTTERAKSQEE 359  
 DB 330 PDAQGEDNSEIKELLEEEIEKESGSEASSEDDEPLPACNGPAQAQSPPTTERAKSQEE 389

QY 360 VLPSTTTPGGALSPGSSSEVTLRTTASGSEOPKKEASTORTSAPSRPPPP 419  
 DB 390 VLPSTTTPGGALSPGSSSEVTLRTTASGSEOPKKEASTORTSAPSRPPPP 449  
 QY 420 RATASPRSSGNIPSSPTASGGGSPTSRASLGTTGTSPTSLEVSNNPPEPKPVRTPE 479  
 DB 450 RATASPRSSGNIPSSPTASGGGSPTSRASLGTTGTSPTSLEVSNNPPEPKPVRTPE 509  
 QY 480 AKENENIHNNQPEELCTSPTLMTSQVASEPGEAKMEDKEKNKLSADSSGQDQLQVS 539  
 DB 510 AKENENIHNNQPEELCTSPTLMTSQVASEPGEAKMEDKEKNKLSADSSGQDQLQVS 569  
 QY 540 MYPENNLLTAPEQEVSTSENPOL 564  
 DB 570 MYPENNLLTAPEQEVSTSENPOL 594  
 RESULT 3  
 ID ABG20324 standard; protein; 594 AA.  
 AC ABG20324;  
 XX  
 DT 13-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #20315.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS  
 XX WO200175067-A2.  
 PN 11-OCT-2001.  
 XX  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS84511.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 50683; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridization probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (I). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 594 AA;  
 Query Match 97.98; Score 2794.5; DB 4; Length 594;  
 Best Local Similarity 98.18; Pred. No. 5e-163; 6; Indels 1; Gaps 1;  
 Matches 554; Conservative 4; Mismatches 4;  
 QY 1 MAEGKAGGAAGLFAKQVKKFSRAQKVLQKLGKAVETKDERFEQASNFYQQAGHKL 60  
 DB 30 MAEGKAGGAAGLFAKQVKKFSRAQKVLQKLGKAVETKDERFEQASNFYQQAGHKL 89  
 QY 61 YKDLKFNLSAVKVMHSSKVSSETIQEISSEWDGHEELKAIWNNDLLWDEYEKLADQ 120  
 DB 90 YKDLKFNLSAVKVMHSSKVSSETIQEISSEWDGHEELKAIWNNDLLWDEYEKLADQ 149  
 QY 121 AVRTWEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAESEFNKA 179  
 DB 150 AVRTWEIYVAQFSEIKERIAKRGKLVYDSARHLEAVQNA-KKDEAKTAKAESEFNKA 209  
 QY 180 QTVFEDLNQELLBELPILYNSRIGCVYTFQNIISNRDVFYREMSKLNHNLYEVMKLEK 239  
 DB 210 QTVFEDLNQELLBELPILYNSRIGCVYTFQNIISNRDVFYREMSKLNHNLYEVMKLEK 269  
 QY 240 QHSNKFVVKGLSSSRSLVISPVPRTATVSSPLTSPSTLSLKSESESVSATEDLA 299  
 DB 270 QHSNKFVVKGLSSSRSLVISPVPRTATVSSPLTSPSTLSLKSESESVSATEDLA 329  
 QY 300 PDAAGEDNSEIKELLEEEIEKEGSEASSEDEPLACNGPAQAPSPPTTERAKSQE 359  
 DB 330 PDAAGEDNSEIKELLEEEIEKEGSEASSEDEPLACNGPAQAPSPPTTERAKSQE 389  
 QY 360 VLPSTTTPGGALSPGSSSEVTLRTTASGSEOPKKEASTORTSAPSRPPPP 419  
 DB 390 VLPSTTTPGGALSPGSSSEVTLRTTASGSEOPKKEASTORTSAPSRPPPP 449  
 QY 420 RATASPRSSGNIPSSPTASGGGSPTSRASLGTTGTSPTSLEVSNNPPEPKPVRTPE 479  
 DB 450 RATASPRSSGNIPSSPTASGGGSPTSRASLGTTGTSPTSLEVSNNPPEPKPVRTPE 509  
 QY 480 AKENENIHNNQPEELCTSPTLMTSQVASEPGEAKMEDKEKNKLSADSSGQDQLQVS 539  
 DB 510 AKENENIHNNQPEELCTSPTLMTSQVASEPGEAKMEDKEKNKLSADSSGQDQLQVS 569  
 QY 540 MYPENNLLTAPEQEVSTSENPOL 564  
 DB 570 MYPENNLLTAPEQEVSTSENPOL 594  
 RESULT 4  
 ID ABG22457 standard; protein; 594 AA.  
 XX  
 AC ABG22457;  
 XX  
 DT 18-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #22448.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS  
 XX WO200175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR

```

PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS86644.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 52816; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (II) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 594 AA;
XX
XX Query Match 97.6%; Score 2787.5; DB 4; Length 594;
XX Best Local Similarity 97.9%; Pred. No. 1.3e-162;
XX Matches 553; Conservative 3; Mismatches 8; Indels 1; Gaps 1;
XX
XX QY 1 MAEGKAGAGAGLFAKOVOKKESRAQKVLQKLGAVETKDERFQSSASNFYQQAAGHKL 60
XX Db 30 MAEGKAGAGAGLFAKOVOKKESRAQKVLQKLGAVETKDERFQSSASNFYQQAAGHKL 69
XX
XX QY 61 YKDLKNFLSAVKVHSESKRYSETLQETIYSSSEWDGHEELKAIYVNNDDLWEDYEELADQ 120
XX Db 90 YKDLKNFLSAVKVHSESKRYSETLQETIYSSSEWDGHEELKAIYVNNDDLWEDYEELADQ 149
XX
XX QY 121 AVRTWEIYVAQSEIKERIAKRGKLVYDSARHLLSAVONA-KDRAKATAKEEFNKA 179
XX Db 150 AVRTWEIYVAQSEIKERIAKRGKLVYDSARHLLSAVONAKDRAKATAKEEFNKA 209
XX
XX QY 180 QTVFEDLNCELLELPILYNSRIGCYVTIFONISNLRDVFYREMSKLNHLYEVMKLEK 239
XX Db 210 QTVFEDLNCELLELPILYNSRIGCYVTIFONISNLRDVFYREMSKLNHLYEVMKLEK 269
XX
XX QY 240 QHSNKFVVKGLSSRRSLVISPPVATVSSPLTSPSTLSLKSESESVATEDLA 299
XX Db 270 QHSNKFVVKGLSSRRSLVISPPVATVSSPLTSPSTLSLKSESESVATEDLA 329
XX
XX QY 300 PDAAGEDNSEIKELLEEEETKEGSEASSSEDDPLPACNGAQAQPSPTTERAKSQE 359
XX Db 330 PDAAGEDNSEIKELLEEEETKEGSEASSSEDDPLPACNGAQAQPSPTTERAKSQE 389
XX
XX QY 360 VLPSSSTTSPGALSFPSCQSSSSATEVVLRTTRTASGEQPKKRASQRTSAPSPPPPP 419
XX Db 390 VLPSSSTTSPGALSFPSCQSSSSATEVVLRTTRTASGEQPKKRASQRTSAPSPPPPP 449
XX
XX QY 420 RATASPRSSGNIPSSPTASGGGSPTRASLGTGTASPTSLSEVPNPEPPKPVRTPE 479

```

---

```

Db 450 RATASPRSSGNIPSSPTASGGGSPTRASLGTGTASPTSLSEVPNPEPPKPVRTPE 509
QY 480 AKENENIHQNPELCTSPILMTSQVASEPGEAKKEDKEKKNKLIADSSSQDQLOYS 539
Db 510 AKENENIHQNPELCTSPILMTSQVASEPGEAKKEDKEKKNKLIADSSSQDQLOYS 569
QY 540 MYPENNNTAPEPOEEVSTSENPOL 564
Db 570 MYPENNNTAPEPOEEVSTSENPOL 594
XX
XX RESULT 5
XX ID ABG21411 standard; protein; 572 AA.
XX AC ABG21411;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #21402.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS85598.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 51770; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 572 AA;

```

AC ABSTRACTS





CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 960 AA;

Query Match 62.6%; Score 1788; DB 4; Length 960;  
 Best Local Similarity 79.7%; Pred. No. 4.8e-101;  
 Matches 369; Conservative 5; Mismatches 5; Indels 84; Gaps 4;  
 QY 46 SASNP--YQQQAEHGKLYKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIV 103  
 DB 144 SAFHFFTYKRAEGHKLYKDLKNFLSAVKVMHSSKRVSETLQFIYSEWYGHLELKAIV 203  
 QY 104 WNDLLWEDYEKLADQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA- 162  
 DB 204 WNDLLWEDYEKLADQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA 263  
 QY 163 KQDEAKTAAEEFNKAQTVFEDLNQELLELPILYNSRIGCVYVTFQNISNLRDVFYRE 222  
 DB 264 KQDEAKTAAEEFNKAQTVFEDLNQELLELPILYNSRIGCVYVTFQNISNLRDVFYRE 323  
 QY 223 MSKLNHLNLYEVMKLEKQHSNKVYVVKGLSSSRSLVSPVRTATVSSPLTSPST 282  
 DB 324 MSKLNHLNLYEVMKLEKQHSNKVYVVKGLS-----SPST 357  
 QY 283 LSLKSESVSATDLAPDAQGEDNSEIKELLEEEIEKEGSEASSSEDDPLPACNGP 342  
 DB 358 LSLKSESVSATDLAPDAQGEDNSEIKELLEEEIEKEGSEASSSEDDPLPACNGP 417  
 QY 343 AQAQPSPTTERRAKSQEVLPSSTTPSPGGALSPSQSPSSATEVVLRTTASEGSEQPKK 402  
 DB 418 AQAQPSPTTERRAKSQEVLPSSTTPSPGGALSPSQSPSSATEVVLRTTASEGSEQPKK 477  
 QY 403 RASIQTSAAPPSPPPPRATASPRSSGNIPSSPTASGGSPSPRASLGCTGTSPTSL 462  
 DB 478 RASIQTSA----- 487

RESULT 9

ABG20323  
 ID ABG20323 standard; protein; 960 AA.

XX AC ABG20323;

XX AC ABG20323;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #20314.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS84510.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 50682; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 960 AA;

Query Match 62.6%; Score 1788; DB 4; Length 960;

Best Local Similarity 79.7%; Pred. No. 4.8e-101;

Matches 369; Conservative 5; Mismatches 5; Indels 84; Gaps 4;

QY 46 SASNP--YQQQAEHGKLYKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIV 103

DB 144 SAFHFFTYKRAEGHKLYKDLKNFLSAVKVMHSSKRVSETLQFIYSEWYGHLELKAIV 203

QY 104 WNDLLWEDYEKLADQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA- 162

DB 204 WNDLLWEDYEKLADQAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA 263

QY 163 KQDEAKTAAEEFNKAQTVFEDLNQELLELPILYNSRIGCVYVTFQNISNLRDVFYRE 222

DB 264 KQDEAKTAAEEFNKAQTVFEDLNQELLELPILYNSRIGCVYVTFQNISNLRDVFYRE 323

QY 223 MSKLNHLNLYEVMKLEKQHSNKVYVVKGLSSSRSLVSPVRTATVSSPLTSPST 282

DB 324 MSKLNHLNLYEVMKLEKQHSNKVYVVKGLS-----SPST 357

QY 283 LSLKSESVSATDLAPDAQGEDNSEIKELLEEEIEKEGSEASSSEDDPLPACNGP 342

DB 358 LSLKSESVSATDLAPDAQGEDNSEIKELLEEEIEKEGSEASSSEDDPLPACNGP 417

QY 343 AQAQPSPTTERRAKSQEVLPSSTTPSPGGALSPSQSPSSATEVVLRTTASEGSEQPKK 402

DB 418 AQAQPSPTTERRAKSQEVLPSSTTPSPGGALSPSQSPSSATEVVLRTTASEGSEQPKK 477

QY 403 RASIQTSAAPPSPPPPRATASPRSSGNIPSSPTASGGSPSPRASLGCTGTSPTSL 462

DB 478 RASIQTSA----- 487

QY 463 EVSPNPPEPKPVRTPEAKENENIHNONPEELCTSPTLMTSQV 505



PA (GEO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 DR GENBANK; O08839.  
 DR  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 588 AA;  
 Query Match 30.18; Score 858; DB 7; Length 588;  
 Best Local Similarity 37.94; Fred. No. 2.8e-44;  
 Matches 230; Conservative 76; Mismatches 179; Indels 122; Gaps 20;  
 QY 1 MAE-GKAGGAGLFAKOVQKFSRAQKVLQKLGKAVETKDERFEGSANSFYQQAEGHK 59  
 DB 1 MAEWSKGVTAGKTASNVQKLTFAQKVLQKLGKADETKDEQFECVQNFNQLTEGTR 60  
 QY 60 LYKDKNPLSAVKVWHESSKRVSETLQRIYSSWDGHEELKAIVNNDLLWEDYBEKLD 119  
 DB 61 LQKDLRTYASVKAMHFAKSLSECLQEVYEPWPGRDEANKTAENNDLLWYHQKLV 120  
 QY 120 QAVRTMIYVAQFEIERIAKGRKLVYDSARHLEAVONA-KDEAKTAK----- 171  
 DB 121 QALLTMTYLGQFPDIKSRIAKGRKLVYDSARHYESLQAKKDEAKIAKPSVLEK 180  
 QY 172 -----AEEFNKAQTVFEDLNQELLELPILYNSRIGCVYT 207  
 DB 181 AAPQWCQKGLQALHVAQTNLLRNQAEELTKAKQVFEEFNVDLQELPSLWNSRVGVYN 240  
 QY 208 IFQNISLRNVFVHEMSKNNHLYEWSKLEKSHSNKVFVVKGLSSS---RSLVISP 264  
 DB 241 TFQIAGLENFHFKEMKLNQNLNDVLVLSLEKOHGNTFTVKAQSDSAPEKGNKSP 300  
 QY 265 VRTATVSPSTPSPTSLKSESSEVSATDELAPDAAGQENSI----- 311  
 DB 301 -----PDGSPAATPEIRVNHPEPPASGASGATIPKSPQLRKGPPVPPPKHT 349  
 QY 312 --KELLLEE-----EIEKEG-----SEASSEED-DPLACNGPAQAQ 346  
 DB 312 --KELLLEE-----EIEKEG-----SEASSEED-DPLACNGPAQAQ 346

Db 350 PSKEMKQEQILSLFDADFVPEISVTTTSPQFAPGPFSEQASLLDLEPLPVPASPVKA- 408  
 QY 347 PSPTTERAKSQEBVLP---SSTTPSPGALSPSGOPSSS--ATEVVLRTTRTASEGSEQPK 401  
 Db 409 PTPSGQS-----IPDLWEPTESQAGVL-PSGEPSSAEGSFVAWPSQTAEPGPAQA 460  
 QY 402 KRASIORTSAPPSPPPPRATASPRSSGNISSPTASCG--GSPTSASLGTG----- 454  
 Db 461 EASEVVGCTGPEGETAASEATSSSLPAVVVTFSTVNGAVGEGSTTTGRLDLPFGFMFKV 520  
 QY 455 -----TASPTSLSEVSNPEPEPEKPEKVTPEAKE-----NENIHNNQPE-ELCTS--P 498  
 Db 521 QAQHDYATDTDELQKAGDVVLVIFQNPESQDEGLMVGKESDWNQHKLEKCRGVFP 580  
 QY 499 TLMTSQV 505  
 Db 581 ENFTERV 587

RESULT 12  
 ADE60840  
 ID ADE60840 standard; protein; 588 AA.  
 XX  
 AC ADE60840;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein O08839, SEQ ID NO 6753.  
 XX  
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX  
 OS Rattus norvegicus.  
 XX  
 FN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GEO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 DR GENBANK; O08839.  
 DR  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 588 AA;

Query Match 30.1%; Score 858; DB 7; Length 588;  
 Best Local Similarity 37.9%; Pred. No. 2.8e-44;  
 Matches 230; Conservative 76; Mismatches 179; Indels 122; Gaps 20;

QY 1 MAE-GKAGAGAGLFAKQVKKFSRAQKVLQKAVETKDERFEGSASNFYQQAGHGH 59  
 DB 1 MAEMSGKGVTAGKIASNVKKLTRAQKVLQKGADETKDEQFCQVQNFNQLTEGTR 60  
 QY 60 LYKDLKNFLSAVKVWHSESKVSETLQEIYSSEWDGHEELKAIWNNDLLWEDYEKLAD 119  
 DB 61 LQKDLRTYLASVKAMHEASKLSECLQEYEPFEGDEANKIAENNDLLWMDYHQLVD 120  
 QY 120 QAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KKDEAKTAK----- 171  
 DB 121 QALLTMDYTLGQFPDKSRIAKGRKLVYDSARHHYVESLQTKAKKDEAKIAPVSLLEK 180  
 QY 172 -----AEEFNKAQVTFEDLNQELLEELELPIIYNRIGCYTT 207  
 DB 181 AAPQWCOQKLAHLVAQTNLLRNQAESELIKAQVFEMNVDLQELPLSNRSGVFFYN 240  
 QY 208 IFQNSLNRLDFVREMSKLNENLYEVMKLEKQHSNKVFVVKGLSSSS---RRSLVISPP 264  
 DB 241 TFSIAGLEENFHKEMSKLNQNLNDVLVLEKQHSNTFVKAQPSDSAPEKGNKSPSP 300  
 QY 265 VRTATVSPISPTSPSTLSKSESVSATEDLAPAAQGENSEI----- 311  
 DB 301 -----PDGSPAATPIRYNHEPEPASGA---SPGATIPKSPQLRKGPPVPPPKHT 349  
 QY 312 ---KELLEEE-----EIEKEG---SEASSSEED-DPLPACNGPAQAO 346  
 DB 350 PSKEMKQQLSLPDDAFVPSIVTSPQFAPGFSEQASLDDLDEPLPVPVAPVKA- 408  
 QY 347 PSPTTERAKSOEVLVLP---SSTTPSGALSPSQSPSS---ATEVVLRTTASEGSEQPK 401  
 DB 409 PTPSQGS-----IPWDLWEPTESQAGVL-PSGEPSSAEGSFAVAMPSTAEFGPAQPA 460  
 QY 402 KEASIQRTSAPPSRPPPRATASPRSSGNTPSSPTASGG--GSPTSPLASLGTG----- 454  
 DB 461 EASEVVGQTQPGETAASEATSSSLPAAVVVTFATVNGAVEGTTTGRDLDPFGMFVK 520  
 QY 455 -----TASPTLSLEVSPPPEPKPVRTPEAKE-----NNINHNQNP-ELCTS--P 498  
 DB 521 QAQHDYATDDELQKAGDVVLVIFQNPQEQDEGLMGVKSQWQHKLEKCKRGVPP 580  
 QY 499 TLMTSQV 505  
 DB 581 ENFTERV 587

RESULT 13  
 ADE60844  
 ID ADE60844 standard; protein; 588 AA.

XX ADE60844;

AC ADE60844;

DT 29-JAN-2004 (first entry)

XX Rat Protein O08839, SEQ ID NO 6757.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN WO2003016475-A2.  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-031247P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX (GEHO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 DR WPI; 2003-268312/26.  
 DR GENBANK; O08839.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 30.1%; Score 858; DB 7; Length 588;  
 Best Local Similarity 37.9%; Pred. No. 2.8e-44;  
 Matches 230; Conservative 76; Mismatches 179; Indels 122; Gaps 20;

QY 1 MAE-GKAGAGAGLFAKQVKKFSRAQKVLQKAVETKDERFEGSASNFYQQAGHGH 59  
 DB 1 MAEMSGKGVTAGKIASNVKKLTRAQKVLQKGADETKDEQFCQVQNFNQLTEGTR 60  
 QY 60 LYKDLKNFLSAVKVWHSESKVSETLQEIYSSEWDGHEELKAIWNNDLLWEDYEKLAD 119  
 DB 61 LQKDLRTYLASVKAMHEASKLSECLQEYEPFEGDEANKIAENNDLLWMDYHQLVD 120  
 QY 120 QAVRTMEIYVAQFSEIKERIAKGRKLVYDSARHLEAVQNA-KKDEAKTAK----- 171  
 DB 121 QALLTMDYTLGQFPDKSRIAKGRKLVYDSARHHYVESLQTKAKKDEAKIAPVSLLEK 180

QY 172 -----ABEENKATQVFDLNOELLELELILYNSRIGCYVT 207  
 Db 181 AAPQWCGKLOAHLVAQTNLLRNQAEELIKAKVFEENVDLOELPSLWNSRVGYVN 240  
 QY 208 IFONISLRDVFVREMSKLNHNLVEVMSKLEKQHSNKVVFVKGSLSSS---RRSLVISPP 264  
 Db 241 TFQSIAGLEENFHKEMSKLNQNLNDVLVLSLEKQHSNTFTVKAQPSDSAPEKGNKSPSP 300  
 QY 265 VRTATVSSPLTPTSPSTLSKSESVSATEDLAPDAAGEDNSEI----- 311  
 Db 301 -----PDGSPAATPEIRVNHEPEPASGA---SPGATIPKSPSLRKGPPVPPPKHT 349  
 QY 312 ---KELLEEE-----EIEKEG---SEASSSED-DLPACNGPAQAQ 346  
 Db 350 PSKEMKQEQILSLFDADFVPEISVTTQFPAEPFSEQASLLDLDFEPLPVASPVKA- 408  
 QY 347 PSPTTERAKSQEVLV-----SSTTPSGGALSPSGQSSS---ATEVLRTRTASEGSPK 401  
 Db 409 PTPSGQS-----IPWDLWEPTESQAGVL-PSGEPSSAEGSFVAVWPSQTAEPPQAQA 460  
 QY 402 KRASIQTSAAPSPRPPPRATASPRSSNIPSSPTASGG---GSPTSPRASLGTG----- 454  
 Db 461 EASEVVGTOEPGCTAASEATSSSLPAVVVETTSATVNGAVEGTTTGRDLDPGFMFVKV 520  
 QY 455 -----TASPRTSLEVPNPEPEKPVRTPEAKE-----NENIHQNPE-ELCTS--P 498  
 Db 521 QAQHDYATDTDELQKAGDVVLVIFQNPPEQDEGLMGVKSQDWNQHKELEKCRGVFP 580  
 QY 499 TLMTSQV 505  
 Db 581 ENFTERV 587

## RESULT 14

ID ADE60852  
 XX ADE60852 standard; protein; 588 AA.

AC ADE60852;

DT 29-JAN-2004 (first entry)

XX Rat Protein O08839, SEQ ID NO 6765.

DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; O08839.

XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 CC injury (CCI) and spared nerve injury (SNI)). In an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 588 AA;

Query Match 30.1%; Score 858; DB 7; Length 588;  
 Best Local Similarity 37.9%; Pred. No. 2.8e-44;  
 Matches 230; Conservative 76; Mismatches 179; Indels 122; Gaps 20;

QY 1 MAE-CKAGGAGLEPAKOVOKKFSRAQKVLQKLGKAVETKDEPESASNFYQQQEGHK 59  
 Db 1 MAEMSGKGVTAGKIASNVQKLTFAQEKVLQKLGKADETKDEQFCVQNFNQLTEGTR 60  
 QY 60 LYKDLKNFLSAVKVMHSSKRVSETLQEIYSSWDGHEELKAIWNNLLWDEYEKLD 119  
 Db 61 LQDLRTVLSAVKAMHEASKLSECLQEVVEPWPGRDEANKIAENNDLLWMDYQKLV 120  
 QY 120 QAVTMEIYVAFSEIEXERIAKGRKLVVDYDSARHLEAVQNA-KKDEAKTAK----- 171  
 Db 121 QALLTMDTYLGQFPDIKSRIAKGRKLVVDYDSARHVEYSLQTAKKDEAKIAPVSLLEK 180  
 QY 172 -----ABEEFNKAQTVFEDLNQELLELPILYNSRIGCYVT 207  
 Db 181 AAPQWCGKLOAHLVAQTNLLRNQAEELIKAKVFEENVDLOELPSLWNSRVGYVN 240  
 QY 208 IFONISLRDVFVREMSKLNHNLVEVMSKLEKQHSNKVVFVKGSLSSS---RRSLVISPP 264  
 Db 241 TFQSIAGLEENFHKEMSKLNQNLNDVLVLSLEKQHSNTFTVKAQPSDSAPEKGNKSPSP 300  
 QY 265 VRTATVSSPLTPTSPSTLSKSESVSATEDLAPDAAGEDNSEI----- 311  
 Db 301 -----PDGSPAATPEIRVNHEPEPASGA---SPGATIPKSPSLRKGPPVPPPKHT 349  
 QY 312 ---KELLEEE-----EIEKEG---SEASSSED-DLPACNGPAQAQ 346  
 Db 350 PSKEMKQEQILSLFDADFVPEISVTTQFPAEPFSEQASLLDLDFEPLPVASPVKA- 408  
 QY 347 PSPTTERAKSQEVLV-----SSTTPSGGALSPSGQSSS---ATEVLRTRTASEGSPK 401  
 Db 409 PTPSGQS-----IPWDLWEPTESQAGVL-PSGEPSSAEGSFVAVWPSQTAEPPQAQA 460  
 QY 402 KRASIQTSAAPSPRPPPRATASPRSSNIPSSPTASGG---GSPTSPRASLGTG----- 454  
 Db 461 EASEVVGTOEPGCTAASEATSSSLPAVVVETTSATVNGAVEGTTTGRDLDPGFMFVKV 520  
 QY 455 -----TASPRTSLEVPNPEPEKPVRTPEAKE-----NENIHQNPE-ELCTS--P 498  
 Db 521 QAQHDYATDTDELQKAGDVVLVIFQNPPEQDEGLMGVKSQDWNQHKELEKCRGVFP 580

QY 499 TLMTSQV 505  
Db 581 ENFTERV 587

RESULT 15

AD660842  
ID ADE60842 standard; protein; 593 AA.

XX ADE60842;

XX 29-JAN-2004 (first entry)

XX Human Protein Q99688, SEQ ID NO 6755.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; Q99688.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 593 AA;

Query Match 29.8%; Score 851.5; DB 7; Length 593;  
Best Local Similarity 36.6%; Pred No. 7e-44;  
Matches 224; Conservative 82; Mismatches 179; Indels 127; Gaps 20;

QY 1 MAE-GKAGAGLFAKQVKKFSAQKVLQKLGKAVETKDRFQSBASNFYQQQAEGHK 59  
Db 1 MAEMSGKVTAGKLTASNVQKLTAKQKVLQKLGKADETKDEQFECQVQNFNKLTEGTR 60  
QY 60 LYKDKNPLSAVKVMHSSKRVSETLQEIYSSSWDGEELKKAIVNNDLLWEDYBEKLD 119  
Db 61 LOKDLRTYLASVXAMHEASKLNCECQEVYEPDWRDEANKIAENNDLLWMDYHOKLVD 120  
QY 120 QAVRTMEIYVAQFSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAK----- 171  
Db 121 QALLTMDTYLQFPDIKSRIAKGRKLVYDYSARHVESLQAKKDEAKIAKPVSLLEK 180  
QY 172 -----ABEEFNKAQTVEDLNQELLEELPILYNSRIGCYVT 207  
Db 181 AAPQWCCQKLAHLVAQTNLRLNQAEEELIKAKVFEEMNVDLQEEPLSLMNSRVGFYN 240  
QY 208 IFQNISLRDVFYREMSKLNHNLYEVMKLEKQHSNKVFVVKGLSSSSRRSLVISPVRT 267  
Db 241 TFQSTAGLEENFHKMSKLNQNLNDVLGLEKHGSGNTFTVKAQPSDN-----APAKGN 294  
QY 268 ATVSPLTSPSTLSLKSESVSATEDLADPAAGGEDNSEI-----K 312  
Db 295 KSPSPDGGSPAATPEIRVNHEPEFAGGA---TPCATLPKSPSQLRGKPPVPPPKHTPK 351  
QY 313 ELLREE-----EIEKEG---SEASSSEED-DLPACNGPAQAQPS 349  
Db 352 EVKQEQILSLFEDTFVPEISVTTTPSQEAPGPFSEQASLLDLDFLPPVTSVKA-PTP 410  
QY 350 TTERAKSQEVLV---SSTTPSGGALSPSGQSSS--ATEVVLRTRTASEGSEQPKKA 404  
Db 411 SGQS-----IPWDLWEPTESAGSL-PSGEPSSAAEGTFVAVSWPQCTAEPGPAQPAEAS 462  
QY 405 SIQRTSAPPRPPPRATASPRSSGNIPS-----SPTASGGSGPTSRAISLGTG 454  
Db 463 EVAGGTQPAQAQEPGFTAASEAASSLPVAVVETPATVNGTVESGSG--AGRLDLPFG 520  
QY 455 -----TASPRTSLEVPNPEPEKPVRTPEAKE-----NENIHNNQNP-ELC 495  
Db 521 FMEKVAQAHDYATDTDBELQKAGDVVLVFPFQNPBEQDEGLMGVKESDWNQKLEKC 580  
QY 496 TS--PTLMTSQV 505  
Db 581 RGVFPENFTERV 592

Search completed: March 4, 2004, 17:26:42

Job time : 75.5656 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:23:55 ; Search time 25.865 Seconds  
(without alignments)  
2097.510 Million cell updates/sec

Title: US-10-069-540A-2  
Perfect score: 2855  
Sequence: 1 MAEGKAGGAGLFAKQYQK.....NNLTAPEQPEVSTENPOL 564

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	829	29.0	482	JC5593	amphiphysin II2 -
2	804.5	28.2	695	S22400	amphiphysin (clone
3	802.5	28.1	682	S22700	amphiphysin - chic
4	463	16.2	461	T22946	hypothetical prote
5	238	8.3	466	T11694	RVS167 protein hom
6	231	8.1	1200	A46194	neurofilament prot
7	230.5	8.1	916	A37864	neurofilament trip
8	228.5	8.0	858	S15752	neurofilament trip
9	214	7.5	763	T08929	hypothetical prote
10	213	7.5	1110	I31116	NF-180 - sea lamp
11	213	7.5	1560	T30282	calcium-binding pr
12	212	7.4	270	T43000	RVS161 protein hom
13	212	7.4	644	S55395	neurofilament prot
14	210	7.4	264	T40661	yeast reduced viab
15	206	7.2	971	T19431	hypothetical prote
16	204.5	7.2	1575	S68448	synaptotagmin, 170K
17	203	7.1	1367	S48478	glucan 1,4-alpha-g
18	202.5	7.1	3507	T34513	hypothetical prote
19	201	7.0	880	D49756	protein T23E7.2b [
20	200.5	7.0	845	A45669	neurofilament trip
21	200.5	7.0	2094	S33124	tpx protein - huma
22	200.5	7.0	4377	A55575	ankyrin 3, long sp
23	200	7.0	1020	QPHUH	neurofilament trip
24	200	7.0	1080	T43164	lacta protein - li
25	197.5	6.9	6642	T29757	protein UNC-89 - C
26	195.5	6.8	2346	T13829	tpx homolog - fru
27	194.5	6.8	706	A45990	junctional sarcopl
28	194.5	6.8	2441	S39161	CREB-binding prote
29	194	6.8	699	T38073	nucleolar phosphop

30	194	6.8	836	2	S49940	cell division cont
31	193.5	6.8	816	2	S64439	hypothetical prote
32	192	6.7	1029	2	T30351	mucin-like protein
33	191.5	6.7	1849	2	C41859	IgA-specific metal
34	191	6.7	742	2	JC7595	scavenger receptor
35	191	6.7	1102	2	T39943	hypothetical prote
36	190	6.7	1664	2	T18262	S-layer protein -
37	189.5	6.6	849	2	S00030	neurofilament trip
38	189	6.6	292	2	S24169	mucin - rat
39	189	6.6	822	2	T51049	related to nucleol
40	189	6.6	870	2	T09137	extensin homolog -
41	188	6.6	473	2	S50755	hypothetical prote
42	188	6.6	1121	2	T02764	myosin-I binding p
43	187.5	6.6	528	2	I47141	gastric mucin (C10
44	187.5	6.6	744	2	F95013	pneumococcal surfa
45	187.5	6.6	2722	2	T20532	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC5593  
amphiphysin II2 - human  
C:Species: Homo sapiens (man)  
C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 09-Jun-2000  
C:Accession: JC5593  
R:Tautsui, K.; Maeda, Y.; Tautsui, K.; Seki, S.; Tokunaga, A.  
Biochem. Biophys. Res. Commun. 236, 178-183, 1997  
A:Title: cDNA cloning of a novel amphiphysin isoform and tissue-specific expression of i  
A:Reference number: JC5593; MUID:97366618; PMID:9223448  
A:Accession: JC5593  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-482 <TSU>  
A:Cross-references: DDBJ:AF001383; NID:G2199534; PIDN:AA61363.1; PID:G2199535  
A:Experimental source: fetal brain  
C:Comment: This protein is involved in the synaptic vesicle recycling and in the regulat  
C:Superfamily: amphiphysin; RVS161 protein homology  
F16-275/Domain: RVS161 protein homology <RVS>  
F1410-481/Domain: SH3 #status predicted <SH3>

Query Match		29.0%	Score 829;	DB 2;	Length 482;
Best Local Similarity		39.2%	Pred. No. 9.9e-29;		
Matches 211;		Conservative	70;	Mismatches 167;	Indels 90;
Gaps		16;			
QY	1	MAE-GKAGGAGLFAKQYQKFSRAQKVLQKLGKAVETKDRFQSASNFYQQAEHGK	59		
Db	1	MAENGSGVITAGTASNVQKLTFAQKVLQKLGKADETKDRFQSCVQNFNKLTEGTR	60		
QY	60	LYKDLKNFLSAVKMHSSKRVSETLQEIYSSWOGHEELKAVNNNDLLWEDYEKLAD	119		
Db	61	LQKDLRTLASVAMEASKGLNECLQEVTEPDWPGDRDANKIAENNDLLMDYQKLVLD	120		
QY	120	QAVTMEIYVAQSEIKERIAKGRKLVYDYSARHLEAVQNA-KKDEAKTAKAEFEFNK	178		
Db	121	QALLTMDTLQGPDIKRSIAKGRKLVYDYSARHYESLQTKKKDEAKTAKAEELIK	180		
QY	179	AQTFEELNQLLEELPILYNSIGCVITFQNIENLRDVFYREMSKLNHNIYEWSKLE	238		
Db	181	AQVFEEENYVDLQEEPLFSLWNSRVGYVNTFQSIAGLEENFHKMSKLNQNLNDLVGLE	240		
QY	239	KQHSNKVFVVKGLSSSSRSRLVSPVTRATVYSSPLTSPSTLSLKSSESVSATEDL	298		
Db	241	KQHSNTFTVYKAQSDN-----APAKGNKSPSPDGSPAATPEVNVHEPPAGGA---	291		
QY	299	APDMAQEDNSEIKELLEEEIEKEGSEASSEEDDPLPACNGPAQAQSPPTTERAKQOE	358		
Db	292	TPGATLPKSPQLRK-----GPPVPPPPKHTPPSEVKQOE	325		
QY	359	EVLP-----SSTTPSGGALSPSQSPSSATEVLRTRTASEGSEKQPKKASTQR	408		
Db	326	QILSLFEDTVPISVITFS-----QP-AAEASEVAGTQTPAA-GAQEPGETRA---	371		

QY 409 TSAPPSPPPRATASPRSSGNIPSSSTASGGSPTSRASLGTG-----TAS 457  
 Db 372 SEASSSLPAVVVEVTFATVNTVTEG-----GSCAG-----RLDLPFGFVKVQAOHDYAT 423  
 QY 458 PRTSLEVPNPPEPKVPTPEAKE-----NENIHQNPE-ELCTS--PTLMTSQV 505  
 Db 424 DTDELQKAGDVVLVIFQNPPEQDEGLMGVKSQWQHELEKCRGVPTFERY 481

RESULT 2  
 S62400  
 amphiphsin (clone 22-2) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 12-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999  
 C:Accession: S62400; 137166  
 R:David, C.; Solimena, M.; de Camilli, P.  
 FEBS Lett. 351, 73-79, 1994  
 A:Title: Autoimmunity in Stiff-Man Syndrome with breast cancer is targeted to the C-term  
 A:Reference number: S48686; MUID:94357284; PMID:8076697  
 A:Accession: S62400  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-695 <DNA>  
 A:Cross-references: EMBL:U07616; NID:G550449; PIDN:AAA21865.1; PID:G550450  
 R:Yamamoto, R.; Li, X.; Winter, S.; Francke, U.; Kilimann, M.W.  
 Hum. Mol. Genet. 4, 265-268, 1995  
 A:Title: Primary structure of human amphiphsin, the dominant autoantigen of paraneoplas  
 A:Reference number: I37166; MUID:95276740; PMID:7757077  
 A:Accession: I37166  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-695 <RES>  
 A:Cross-references: EMBL:X81438; NID:G662991; PIDN:CAA57197.1; PID:G662992  
 C:Genetics:  
 A:Gene: GDB:AMPH  
 A:Cross-references: GDB:386990  
 A:Map position: 7p14-7p13  
 C:Superfamily: amphiphsin; RVS161 protein homology  
 F:11-270/Domain: RVS161 protein homology <RVS>

Query Match 28.2%; Score 804.5; DB 2; Length 695;  
 Best Local Similarity 34.3%; Pred. No. 1.7e-27;  
 Matches 206; Conservative 104; Mismatches 175; Indels 115; Gaps 18;

QY 11 GLFAQVQKFRAGAEKVLQKLGKAVETKDERPQSASNFYQQAEGHLYKDLKNPLSA 70  
 Db 7 GFAKNVQKRLNRAQAEKVLQKLGKADETKDEQFEYVQNFKRQAEGRLORELRGLAA 66  
 QY 71 VKVMHSSKRVSETLQEIYSSSEWDGHEELKAIWNNNDLLMEDYEELADQAVRTMELYA 130  
 Db 67 IKMQEASMKLTESLHEVTEPDWYGRDVKWVGKCDVWDFHQLVDGSLTLDLYLG 126  
 QY 131 QFSEIKERIAKGRKLVYDSARHLEAVQNAK-KDEAKTAKAEFEFNKAQTVFEDLNQE 189  
 Db 127 QPFDIKNRIAKRSKLVYDSARHLEALQSSKRKESRISKAEFEFQKAKVFEFNVD 186  
 QY 190 LLEELPILYNSRIGCVYIFQNI SNLRDVFREMSKLNHLYEWSKLEKHQNKVFPVK 249  
 Db 187 LOEELPILYNSRIGCVYIFQNI SNLRDVFREMSKLNHLYEWSKLEKHQNKVFPVK 246  
 QY 250 GLSSSS---RRSLVSPVVRATVSPSLTS-----PTSPSTLSKSESVSATDLAP 300  
 Db 247 GAPSDGSLPIAKTSPPEEPLSPSTASPNHILAPAPAPRPSQT----- 297  
 QY 301 DRAQGEDNSEIKELLEBEIEKEGSEASSEDPLP--ACNGPAQAPSPPTTERAKQE 358  
 Db 298 --RKGPVPPLPKVFTTKELQOE--NIISFFEDNFVPEISVTPSQNE---VPEVKKE 349  
 QY 359 EYL-----PSSTTPSPGALSPGSPSS-----SATEVLRRTASEGS-----EOP 400  
 Db 350 TLILDLDLDFPKFVTPASAGVTHSPMSQTLPWDLWTITSLDV---QFASGGSFNGFTQP 406

QY 401 KKRASI-----QRTSAPPSPPPRATASPRSSGNIPSSSTASGGSPTS 447  
 Db 407 QDTSLFTWQTDQSMICNLAESQAPPTPEKAEPLAAVTPAVG----- 449  
 QY 448 RASLGTGTASPRTSLEVPNPPEPKVPTPEAKEENIHQNPEELCTSPTLMTSQVAS 507  
 Db 450 -LDLGMOTRA-----EETVEAVIIPGADADAAGCT-----LVSAAG 486  
 QY 508 EPG---EAKM-----EDKDKONKLISADSSGQDQLQVSMVFPENNLTAPBPQEVST 559  
 Db 487 APGEAEAEKATVPAGEGVSLSEAKIGTETTEGAESAQ-----PEAELEATVPQEKVIPS 542

RESULT 3  
 S22700  
 amphiphsin - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: S22700  
 R:Lichte, B.; Voh, R.W.; Meyer, H.E.; Kilimann, M.W.  
 EMBO J. 11, 2521-2530, 1992  
 A:Title: Amphiphsin, a novel protein associated with synaptic vesicles.  
 A:Reference number: S22700; MUID:92331604; PMID:1628617  
 A:Accession: S22700  
 A:Molecule type: mRNA  
 A:Residues: 1-682 <LIC>  
 A:Cross-references: EMBL:X60422; NID:G62842; PIDN:CAA42953.1; PID:G62843  
 C:Superfamily: amphiphsin; RVS161 protein homology  
 F:11-270/Domain: RVS161 protein homology <RVS>

Query Match 28.1%; Score 802.5; DB 1; Length 682;  
 Best Local Similarity 33.1%; Pred. No. 2e-27;  
 Matches 206; Conservative 118; Mismatches 203; Indels 95; Gaps 19;

QY 1 MAEGKAGAGLFAKQVQKFRAGAEKVLQKLGKAVETKDERPQSASNFYQQAEGHLYK 60  
 Db 1 MADNK---TGIFAKNVQKRLNRAQAEKVLQKLGKADETKDEQFEYVQNFKRQAEGRSL 56  
 QY 61 YKDLNPLSAVKVMHSSKRVSETLQEIYSSSEWDGHEELKAIWNNNDLLMEDYEELADQ 120  
 Db 57 QRELRAVLAAIKGMQDASKKLTSLHEVTEPDWYGRDVKWVGKCDVWDFHQLVDG 116  
 QY 121 AVRTMELYVAQFSEIKERIAKGRKLVYDSARHLEAVQNAK-KDEAKTAKAEFEFNKA 179  
 Db 117 SLTLDLYLQGFDPDKIRIAKRSKLVYDSARHLEALQSSKRKDEGRITAKAEFEFQKA 176  
 QY 180 QTFEDLNQELLELPILYNSRIGCVYIFQNI SNLRDVFREMSKLNHLYEWSKLEK 239  
 Db 177 QKVEEFTDLOELPILYNSRIGCVYIFQNI SNLRDVFREMSKLNHLYEWSKLEK 236  
 QY 240 QHSNKVFWVVGKLSSSS---RRSLVSPVVRATVSPSLTS-----PTSPSTLSKSES 290  
 Db 237 QHADKAFITQCAPSDGSLPIAKTSPPEEPLSPSTASPNHILAPAPAPRPSQT 296  
 QY 291 SVSATEDLAPDAAGEDNSEIKELLEBEIEKEGSEASSEDPLPACNGPAQAPSP-P 349  
 Db 297 -----LRKGPVPPLPKVFTTKELQOE--NIISFFEDNFVPEISVTPSQNEIP 343  
 QY 350 TTERAKS-----QEEVLPSTTPSPGALSPSG-----QPSS----- 382  
 Db 344 ETKKVESLLDLDLDFPKFVTPASAGVTHSPMSQTLPWDLWTITSELVQSPASFAFGPAQD 403  
 QY 383 ATEVLR-----TETASEGSEQPKKASQRTS-----APPSPPPPPRA---TA 423  
 Db 404 TPAFVGSNNVNTETLTAEEAFGLBELKVEETPAAVVEKEAILAEFPDFTQAESEIEA 463  
 QY 424 SPRSSG--NIPSSSTASGGSS---PTSFRASLGTGTASPRTSLEVPNPPEPKVPT 478  
 Db 464 GDKETTGAIEKSEVSVSAGAVAVEDSVVVAAGAGGAVRTEQEAAGADKPGQEKDV 523  
 QY 479 EAKENIHQNPEELCTSPTLMTSQVASPEPGKQWEDKDKONKLISAD--SSGQDQLQ 537  
 Db 524 DVSQ-----EKVSSIPSVVIEPASNNEGEEHHVIMNESKDAAEAMTQGTDS--E 573



QY 538 VSMVPENNLTAPEQREVST 559  
 Db 574 TSQIGSEQAT-----EIIQT 590

## RESULT 4

T22946  
 Hypothetical protein F58G6.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000  
 C:Accession: T22946

R.Lloyd, C.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19641

A:Accession: T22946

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-461 <WIL>

A:Cross-references: EMBL:Z68217; PIDN:CAA92465.1; GSPDB:GN00022; CESP:F58G6.1

A:Experimental source: clone F58G6

C:Genetics:

A:Gene: CESP:F58G6.1

A:Map position: 4

A:Introns: 19/3; 57/3; 96/3; 130/3; 157/3; 186/3; 219/3; 271/3; 302/3; 335/3; 357/3

C:Superfamily: amphiphysin; RV5161 protein homology

Query Match 16.2%; Score 463; DB 2; Length 461;  
 Best Local Similarity 26.9%; Pred. No. 4.1e-13;  
 Matches 144; Conservative 96; Mismatches 175; Indels 120; Gaps 18;

QY 10 AGLPAKOVKFSRAQKVLQKLGKAVETKDERFEQASNFYQQAEGHLYKDLKNFLS 69

Db 2 ADLFNKILKATNTKELLEGIGKATQDEVFDQAAANLNKQSKCEKLHKVKNYSS 61

QY 70 AVKVMHSSKRVSETLOEIIYSSEWDGHEELKATVWNNLLWEDYKELADQAVRTWEIV 129

Db 62 ALRTLLSAEQRLRTIDAVEPEWPDREHLTAIFDNLIDIQTELEKTVCDLPOIVTQV 121

QY 130 AOFSEIKERAKRGLVDYDSARHLEAVQ-NAKQ-EAKTAKABEEFNKAQTVFEDLN 187

Db 122 NQFDLKKKIEKGRKLVYDSAKNSFNVS VKASSKNDNPFLAKATMELQAAEQMYTEM 181

QY 188 QELLEELPILYNGRIGCVYVIFONISNLRDVFVREMSKLNHLYEVMVSKLEKOHNSKVFV 247

Db 182 NELLEILPAVFDSEITFTVDTLQTLFNANSVYQTDASKFHK---QIVMQLDK--- 230

QY 248 VKGLSSRSRLVISPVRATV--SSPLTSPSTLSKSESVSATEDLAPDAAQ 305

Db 231 ---LGESMDYLRVARPEARALTPTDITSLASSDAPAT-----PDQNK- 269

QY 306 EDNSEIKELLEEEIEKEGSEASSSEDDPLPACNGPAQAPSPITERAKSQEEVLPSST 365

Db 270 SDNSSLRQ-----TPSPS-----VLPS-- 286

QY 366 TPSPGALSPSGOPSSSATEVILRTTASGSEQPKRASIORTSAPSPRPPPRATASP 425

Db 287 -PVP-----APWAKPRESVSMANSTNPFDDDDAEHDKAT-----PTDVEEKEFKYVP 335

QY 426 RPSSGNIPSPSTASGGSPSPASISLTGTASP-----RTSLEVSP---NPEPEKPV 475

Db 336 KLNAAQTAQAQAIAAKREKKSAS-----NPFDDDEDESTEVEAPKCKDPEPTPKPL 389

QY 476 RTPAEKENENHNQNEEL--CTS-----PTLMTSQVASPGBAKQMKDKKNKL 524

Db 390 -----DGIETKEKRTLVFTVSTNYKAVDTDELSPFGQVQIKVIEANEDDQL 436

## RESULT 5

T11684

RV5167 protein homolog - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000

C:Accession: T11684

R.Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z17313

A:Accession: T11684

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-466 <SEE>

A:Cross-references: EMBL:AL031536

A:Experimental source: strain 972h(-)

C:Genetics:

A:Map position: IIR

A:Introns: 444/3

A:Note: SPSC21D10.12

C:Superfamily: RV5167 protein; RV5161 protein homology; SH3 homology

F:4-270/Domain; RV5161 protein homology <RVS>

F:414-463/Domain; SH3 homology <SH3>

Query Match 8.3%; Score 238; DB 2; Length 466;  
 Best Local Similarity 22.0%; Pred. No. 0.0017;  
 Matches 110; Conservative 69; Mismatches 194; Indels 128; Gaps 16;

QY 15 KQVQKFSRAQKVLQKLGKAVETKDERFEQASNFYQQAEGHLYKDLKNFLSAVKVM 74

Db 4 KGFTKALARTPQTLRSKFNVEITKDPFIYEDAGRRFKSLETEAKLAEDAKKYTDALNG 63

QY 75 HESSKRVSETLOEIIYSSEWDGHEELKATVWNNLLWEDYEEK 116

Db 64 LNHQIGFADACIIYKIPISGRASDPESYEQEAGNAGIEAAEAYKEIVY-----DLQKN 116

QY 117 LADQAVRTWEIVYVQF-----SEIKERIAKRGKLVYDSARHLEAVQNAK-- 163

Db 117 LASE-----MDVINTRIVNPTGELLKIVKDVDKLLKRDHKQLDYDRHRSFCKLOEKDK 172

QY 164 --KDEAKTAKABEEFNKAQTVFEDLNQELLEPLILYNSRIGCVYVIFONISNLRDVFVR 221

Db 173 SLQDEKGLYEATAPFQSSQEQEYVYNEMLKEELPKLFAQAQFIAPLFG-----FY 225

QY 222 EMKLNHNIYEVMSKLEKOHNSKVFVVKGLSSRSRLVISPVRATVTSPLTSPSPS 281

Db 226 MQLNVYVLYYKRMHCEIQYFD--FNTDILESYYERRGDKVDRAEALITTKFTAKPTYK 283

QY 282 TLLKSESSESVSATEDLAPDAAQEDNSEIKELLEEEIEKEGSEASSSEDDPLPACNG 341

Db 284 RPNQGGCGKDATAS-----SSSFSSSKREE----- 308

QY 342 PAQAQSPPTTERAKSQEEVLPSSTTPSPGALSPSGOPSSSATEVILRTTASGSEQPK 401

Db 309 -AAAEPSST-----ATDIPPPYSTPSVAGASDYS--TPSAGYQTVQTTTTTTEAAAAQY 360

QY 402 KRASIQRTSAPSPRPPPRATASPPSGNIPSPSTASGGSPSPASISLTGTASPTSPS 461

Db 361 PQAAF-----PPPP--VMPQFAA-----AAVTPVAAPVAA 389

QY 462 LEVS-PNPEPPEKPVRTPEAK 481

Db 390 AAAAVPVPFPAPAPAAAPAE 410

## RESULT 6

A46194

neurofilament protein NF-220, high-molecular-weight splice form - longfin squid

C:Species: Loligo pealeii (longfin squid)

C>Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 17-Mar-2000

C:Accession: A46194

R;Way, J.; Hellmich, M.R.; Jaffe, H.; Szaro, B.; Pant, H.C.; Gainer, H.; Battey, J.

Proc. Natl. Acad. Sci. U.S.A. 89, 6963-6967, 1992

A:Title: A high-molecular-weight squid neurofilament protein contains a lamin-like rod d

A:Reference number: A46194; MUID:92357751; PMID:1379729

A:Accession: A46194

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1200 <WAY>

A;Cross-references: GB:M94389; NID:g161291; PID:g161292  
 A;Experimental source: stellate ganglion  
 A;Note: sequence extracted from NCBI backbone (NCBIP:113499)  
 C;Superfamily: neurofilament triplet H protein  
 C;Keywords: alternative splicing

Query Match 8.1%; Score 231; DB 2; Length 1200;  
 Best Local Similarity 18.3%; Pred. No. 0.011;  
 Matches 132; Conservative 129; Mismatches 240; Indels 222; Gaps 24;  
 QY 18 OKKPSRAQKV-----LQKLGAVTVDKPEGASNFYQQAGHLYKDLNKLFSAV 71  
 DB 196 QGCIARQQLADLEGEISMLRSLSELEKMKQSNILAKQNDMEKMDLNN-----E 251  
 QY 72 KWHSESSKVSSTLQEIYSSSDGH-BELKAIV-----WNDL-----LW 110  
 DB 252 TINHLDAENRRGTLEEELEFQKVHQAQLKELALAYDITTAENRFRWELAQARDIQ 311  
 QY 111 EDEYKLAQDAQVITMEIY-----VAQF-----SEKIERIAKGRKLVD 148  
 DB 312 QEYDAK-CDQMRGDIAYYNLKVQEFRTGATKQNNVTRNKEENTKLKSNNTBI RNKLD 370  
 QY 149 YDSARHLEAV-----QNAKQDAKTAKEEFNKATVPEDLNQELLEELPILYN 199  
 DB 371 LEARNAQLERTNQDLRLQEKDQNELESQYKEEITKLRGENESILKELQDMDIKUS 430  
 QY 200 SRIGCVITFQINSLRDVFRYREMSKLNHLYEYMSKLEKQHNKVFVVKGLSS----- 254  
 DB 431 LEL-----EIAAYRKLLEGEESRIELVHPMTIGTREAYRPELITKNGKASDDSS 482  
 QY 255 ---SRSLVIPPVPTATVSSPLTS-----PTSPSTLSKSESVSATDEL-- 298  
 DB 483 KDGTVRAKSPVDVVAETKLTITSYCGGDADDEGKDSDDTHTEAEABETRADSDAT 542  
 QY 299 -----APDAAQEDNSEIKELLEEEIEKEGSEASSEEDD 334  
 DB 543 GTGLDEVKEESVLKSEKDKSVKRDDEEEDNDQTESEVDEDEADIEKSKSESDKTE 602  
 QY 335 PLPCNGAQAQAPPT-----TERAKSQEVLPSST-----PFGG----- 371  
 DB 603 SDKAESERAEVKATTPESDKAESVKSRSSTTAFIDKPKPGKRDTLFQSDKV 662  
 QY 372 -----ALSP-----SQQSSSATEVLRTRA 393  
 DB 663 TSPVLLERAMSLQAPTEKPLSDTIKSPVVSFAISFVSVRGATSTPAERTSPISGG 722  
 QY 394 SEGSEQPKKRASIQRTSAPPR-----PPPRATASPPSSGNIPSSPTASGGSP 444  
 DB 723 SEKSAKSPVRS--EATKSPVSEKAGSKSPVPSERALSPLVSESVHSTAMSHTSRSP 780  
 QY 445 TSPRASLTGTASPTSLVSPNPEPP--EKPVRTPEAKENIHNQNPBELCTSTPLMT 502  
 DB 781 TASEKSVKSPHSERTA--SPTAKSPIIPEAKSPKDESEKLSPESESVMGSKSQIT 837  
 QY 503 SQVASEPGEAKMEDKEKNKLISADSEGGDQLQVSWPENNLTAPPEQBEV---STS 559  
 DB 838 SSSAKSPVPSBK-----ADSEK-----SATSPTESEKVDDESSA 871  
 QY 560 ENP 562  
 DB 872 RSP 874

RESULT 7  
 A27864  
 neurofilament triplet M protein - human  
 N;Alternate names: NF-M (medium) protein  
 C;Species: Homo sapiens (man)  
 C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 13-Aug-1999  
 C;Accession: A27864; A30157  
 R;Myers, M.W.; Lazzarini, R.A.; Lee, V.M.Y.; Schlaepfer, W.W.; Nelson, D.L.  
 EMBO J. 6, 1617-1626, 1987  
 A;Title: The human mid-size neurofilament subunit: a repeated protein sequence and the

A;Reference number: A27864; MUID:87275853; PMID:3608989  
 A;Accession: A27864  
 A;Molecule type: DNA  
 A;Residues: 1-916 <MYE>  
 A;Cross-references: GB:Y00067; NID:g35045; PID:CAA6276.1; PID:g35046  
 R;Lee, V.M.Y.; Otvos Jr., L.; Carden, M.J.; Hollosi, M.; Dietzschold, B.; Lazzarini, R.A.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 1998-2002, 1988  
 A;Title: Identification of the major multiphosphorylation site in mammalian neurofilament  
 A;Reference number: A30157; MUID:88158120; PMID:2450354  
 A;Contents: annotation; phosphorylation sites  
 C;Genetics:  
 A;Introns: 360/3; 402/2  
 C;Superfamily: cytoskeletal keratin  
 C;Keywords: coiled coil; phosphoprotein  
 F;615,628,641,654,680/Binding site: phosphate (Ser) (covalent) #status predicted  
 P;615,628,641,654,680/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 8.1%; Score 230.5; DB 2; Length 916;  
 Best Local Similarity 22.4%; Pred. No. 0.0081;  
 Matches 131; Conservative 96; Mismatches 264; Indels 93; Gaps 24;  
 QY 41 ERPEQASNFYQQAGHLYKDLNKLFSNAVKMHSESKRVSETLQEIYSSSEWDGHEELK 100  
 DB 187 ERFEERARLDDTEAIRAURKDIIE-ASLVKV--ELDKKV-OSLQDEVAFLRSHEEV 242  
 QY 101 AIVWNNDLLWE-----DYEKLADQAVRTMEIYVAQFS-----EIKERIAKEGR 144  
 DB 243 A-----DLAQIQASHITVERKDYLTDTSTALKEIRSQLESHSDQNMHQAEEFKCVA 297  
 QY 145 KLVDYDSARHLLEAVQNAKDEAKTAAEE---EPNKAQTVFDLNQELLEELPILYN 200  
 DB 298 KLT--EAEQNKRAISAKKEEIAEYRQLOQSKSIELESVGTGKESLRQ-LSDIERHNH 354  
 QY 201 RIGCVITFQINSLRDVFRYREMSKLNHLYEYMSKL-EKQHSNKVFVVKGLSSSSRS 259  
 DB 355 DLSSYQDTIQLEN-----ELRGTKWEMARHLREYQDLLNVKMDIEIAAYRKL 404  
 QY 260 VIQPPVTRATVSPPLTSP--TSPSTLSKSESVSVA-----TEDLAPDAAGE 306  
 DB 405 LEGETRFSTFAGSIITGPLYTHRPITITISSKTQKTVEAPKLKVQKFEVEIIEETKVD 464  
 QY 307 DNESEIKELL---EE-----EEIEKEGSEASSSE---EDDPLPACNGPAQA-QPSPTTER 353  
 DB 465 EKSEMEALTAITEELAAASKKEKKAEEKEEPEAEHEEVAAKSPVKATAPVKEEE 524  
 QY 354 AKSQEVLPSSTTPPGGALSPGQSSSATVWL-----RTRTASEGSE-OPKK 402  
 DB 525 GEKEESEGQEEEEEDEGAKSDQABEGSGEKGSEKEGEQEGETEAEAEGEAEAKE 584  
 QY 403 RASIQRTSAPPSPRPPTA---SPRPSSGNIPSPSTASGGSP--SPRASLTGTAS 457  
 DB 585 EKKVEKSEEVATKEELVADAKVPEKAKSPVKSPVVEEKGSVPVKSPVVEEKGSVPV- 643  
 QY 458 PRTSLVSPNPEPPEKPVRTPEAKENIHNQNPBELCTSP-----LMTSOVASEPGEAK 513  
 DB 644 PKSPVVEEKGSVPVKSPV--BEKKGSPVKSPVVEEKAKSPVKSPVVEEAKSKAEVCKGE 700  
 QY 514 KMEDEKDKNKLISADSEGGDQLQVSWPENNLTAPPEQBEVS 557  
 DB 701 QKEEKEKVEKPAKPEKKEKPKDVPKPKAESPVKKEAVA 744

RESULT 8  
 S15762  
 neurofilament triplet M protein - chicken  
 C;Species: Gallus gallus (chicken)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 13-Aug-1999  
 C;Accession: S15762; S08061; A27040  
 R;Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.D.  
 Nucleic Acids Res. 18, 521-529, 1990  
 A;Title: Isolation of the chicken middle-molecular weight neurofilament (NF-M) gene and  
 A;Reference number: S15762; MUID:90174973; PMID:2106668  
 A;Accession: S15762  
 A;Status: preliminary





QY 474 PVRTPEAKENENIHNONPEELC--TSPTLMTQVASEPGCAKMKDEKKNKLISADS-S 530  
Db 455 PRDVPKKAESFWKEAEAEAAATIKPTVGLKEKTEKEKPLQKEKEKAGEEGGSEE 514  
QY 531 EGQDQ-----LQVSMVPENNLLTAPEPOREVSTSE 560  
Db 515 EGSDDQSKRAKKEDIAVNGEGEGKEEEETPEKKGSGRE 553

RESULT 14  
T40661  
yeast reduced viability upon starvation protein 161 homolog, implicated in cell growth a  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
C:Accession: T40661  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.  
submitted to the EMBL Data Library, November 1998  
A:Reference number: Z21889  
A:Accession: T40661  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-284 <LYN>  
A:Cross-references: EMBL:AL034352; PIDN:CAA22181.1; GSPDB:GN00067; SPDB:SPBC725.09c  
A:Experimental source: strain 972h-; cosmid C725  
C:Genetics:  
A:Gene: SPDB:SPBC725.09c  
A:Map position: 2  
A:Introns: 16/3  
C:Superfamily: RVS161 protein; RVS161 protein homology  
F:4-255/Domain: RVS161 protein homology <RVS>  
Query Match 7.4%; Score 210; DB 2; Length 264;  
Best Local Similarity 26.6%; Pred. No. 0.014;  
Matches 63; Conservative 44; Mismatches 110; Indels 20; Gaps 6;

```

Db      7 KCAVNRAGTSVMKTKTGHVETVDRFBERRVMTMESAAKKLQKEAGYLDALRAMTAS 66
QY      78 SKRVSEITLQRIYSEWDGHEELKAIWVNDLLWEDYEE---KLDAQVRTMEI-----YV 129
Db      67 QTRANTIDAFYGA--GSKDGVSAIYRQ--VVEDLDADTVKELDGPRTTVLDPSISRFC 122
QY      130 AQFSEIKERIAKRGLKVYDYSARHLE-AVQNQAKDDEAKTAKAESEFNKAQTVPEDLNQ 188
Db      123 SYFPDINAAITKENHKLLDHDAMRAKVQKLVDKPSNDITKLPRTKEAAMAEVYETLNN 182
QY      189 ELLEELPILYNSRIGCVVITFIQNISLNRDVFVREMSKLNHLNLYEVMWSKLEKHSNKV 245
Db      183 QLVSLEPQLTALRPVLDPSFEALVKIQRFCREG-----YEKAAQVOQYFDSNV 232

RESULT 15
Ti19431
hypochemical protein C25A1.10 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C.Accession: T19431
R.Mortimore, B.
submitted to the EMBL Data Library, October 1996
A.Reference number: Z19124
A.Accession: T19431
A.Status: preliminary; translated from GB/EMBL/DBDJ
A.Molecule type: DNA
A.Residues: 1-971 <N1>
A.A.Cross-references: EMBL:Z81038; PIN: CAB02755.1; GSPDB:GN00019; CESP:C25A1.10
A.Experimental source: clone C25A1
C.Genetics:
A.Gene: CESP:C25A1.10
A.Map position: 1
A.Introns: 38/3; 92/3; 201/3; 919/3
A.Superfamily: nucleolus-cytoplasm shuttle phosphotransferase

```

Search completed: March 4, 2004, 17:31:42  
Job time : 28.865 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:21:39 ; Search time 14.3694 Seconds  
(without alignments)  
2043.754 Million cell updates/sec

Title: US-10-069-540A-2

Perfect score: 2855

Sequence: 1 MARGKAGGAGLFAKQVQKK.....NNLTAPQEVSVISENPQL 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	858	30.1	588	1 BIN1_RAT	O08839 rattus norv
2	851.5	29.8	593	1 BIN1_HUMAN	O00499 homo sapien
3	847	29.7	588	1 BIN1_MOUSE	O08539 mus musculu
4	815.5	28.6	683	1 AMPH_RAT	O08838 rattus norv
5	804.5	28.2	695	1 AMPH_HUMAN	P49418 homo sapien
6	802.5	28.1	682	1 AMPH_CHICK	P50478 gallus gall
7	238	8.3	466	1 HOBA1_SCHPO	O74352 schizosacch
8	232	8.1	810	1 NFM_BOVIN	O77788 bos taurus
9	230.5	8.1	915	1 NFM_HUMAN	P07197 homo sapien
10	228.5	8.0	857	1 NFM_CHICK	P16053 gallus gall
11	212.5	7.4	2517	1 NCR2_HUMAN	O9y618 h nuclear r
12	212	7.4	644	1 NFM_RABIT	P54938 oryctolagus
13	210	7.4	264	1 HOB3_SCHPO	O9um77 schizosacch
14	208.5	7.3	1090	1 ASP1_HUMAN	O96kq4 homo sapien
15	204.5	7.2	1324	1 SYJ1_BOVIN	O18964 bos taurus
16	204.5	7.2	1574	1 SYJ1_RAT	O62910 rattus norv
17	203	7.1	1367	1 AMPH_YEAST	P08640 saccharomyc
18	201.5	7.1	1087	1 ASP1_MOUSE	O62415 mus musculu
19	200.5	7.0	2349	1 TPR_HUMAN	P12270 homo sapien
20	200.5	7.0	4377	1 ANK3_HUMAN	O12955 homo sapien
21	200	7.0	845	1 NFM_EAT	P12839 rattus norv
22	199.5	7.0	518	1 TBM4_DROME	P49455 drosophila
23	199.5	7.0	622	1 BPB1_HUMAN	O9y313 homo sapien
24	199	7.0	1026	1 NFM_HUMAN	P12036 homo sapien
25	198	6.9	1338	1 ACIN_MOUSE	O9jix8 mus musculu
26	197.5	6.9	6632	1 UN89_CAREL	O01761 caenorhabdi
27	197	6.9	1360	1 TNIK_HUMAN	O9uke5 homo sapien
28	194.5	6.8	705	1 TRDN_RABIT	O28820 oryctolagus
29	194.5	6.8	2441	1 CBP_MOUSE	P45481 mus musculu
30	194	6.8	699	1 NP14_HUMAN	O14978 homo sapien
31	194	6.8	836	1 NOT3_YEAST	P06102 saccharomyc
32	194	6.8	1575	1 SVJ1_HUMAN	O43426 homo sapien
33	194	6.8	3644	1 MINT_MOUSE	O62504 mus musculu

34	193.5	6.8	816	1 YG3A_YEAST	P53278 saccharomyc
35	193.5	6.8	5085	1 PCLO_RAT	O9jks6 rattus norv
36	193	6.8	758	1 K179_HUMAN	Q14684 homo sapien
37	191.5	6.7	1849	1 IGA4_HAEIN	P45386 haemophilus
38	191	6.7	1102	1 YG49_SCHPO	O60184 schizosacch
39	190.5	6.7	3664	1 MINT_HUMAN	O96t58 homo sapien
40	190	6.7	1864	1 SLP1_CLOTH	Q08852 clostidium
41	189.5	6.6	848	1 NFM_MOUSE	P08553 mus musculu
42	187	6.5	700	1 TRDN_CANFA	P82179 canis famli
43	187	6.5	1087	1 NFM_MOUSE	P19246 mus musculu
44	187	6.5	1248	1 DIA1_HUMAN	O60610 homo sapien
45	186.5	6.5	1694	1 IGA0_HAEIN	P44969 haemophilus

ALIGNMENTS

RESULT 1

BIN1\_RAT

ID BIN1\_RAT STANDARD; PRT; 588 AA.

AC O08839;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Myc box dependent interacting protein 1 (Bridging integrator 1)

DE (Amphiphysin-like protein) (Amphiphysin II).

GN BIN1 OR AMPHL OR AMPH2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

LN [1]

RP SEQUENCE FROM N.A., SUBUNIT. AND ALTERNATIVE SPLICING.

RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex, and Kidney;

RX MEDLINE=9809145; PubMed=9348539;

RA Wigge P., Kohler K., Hunt S., Doyle C., McMahon H.T.;

RT "Amphiphysin heterodimers: potential role in clathrin-mediated endocytosis."

RL Mol. Biol. Cell 8:2003-2015(1997).

LN [2]

RP SEQUENCE FROM N.A. (ISOFORM AMPH2-1).

RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex;

RX MEDLINE=97424383; PubMed=9280305;

RA McMahon H.T., Wigge P., Smith C.;

RT "Clathrin interacts specifically with amphiphysin and is displaced by dynamin."

RL PNAS Lett. 413:319-322(1997).

LN [3]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 506-588, FUNCTION, AND INTERACTION WITH DNMI.

RX MEDLINE=98409533; PubMed=9736607;

RA Owen D.J., Wigge P., Vallis Y., Moore J.D.A., Evans P.R., McMahon H.T.;

RT "Crystal structure of the amphiphysin-2 SH3 domain and its role in the prevention of dynamin ring formation."

RL EMBO J. 17:5273-5285(1998).

CC -!- FUNCTION: May be involved in regulation of synaptic vesicle endocytosis. May act as a tumor suppressor and inhibits malignant cell transformation.

CC -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).

CC Binds to SYN1 and DNMI through its SH3 domain, and to clathrin through a region outside of the SH3 domain. Also binds to alpha-adaptin. Interacts with the N-terminal transactivation domain of MYC in a manner requiring the integrity of the conserved MYC box regions 1 and 2.

CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=6;

CC Name=AMPH2-1;

CC IsoId=O08839-1; Sequence=Displayed;

CC Name=AMPH2-2;

CC IsoId=O08839-2; Sequence=VSP\_000260;

CC Name=AMPH2-3;





RT suppressor.";  
RL Nat. Genet. 14:69-76(1996).  
RN [4]  
RP REVISIONS TO N-TERMINUS.  
RA Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS IIB; IIC1; IIC2 AND IID).  
RC TISSUE=Brain;  
RX MEDLINE=98264340; PubMed=9603201;  
RA Ramjaun A.R., McPherson P.S.;  
RL "Multiple amphiphysin II splice variants display differential clathrin  
RT binding; identification of two distinct clathrin-binding sites.";  
RL J. Neurochem. 70:2369-2376(1998).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORMS IIB AND IIC3).  
RC TISSUE=Brain;  
RX MEDLINE=97366618; PubMed=9223448;  
RA Tautsui K., Maeda Y., Tautsui K., Seki S., Tokunaga A.;  
RL "cDNA cloning of a novel amphiphysin isoform and tissue-specific  
RT expression of its multiple splice variants.";  
RL Biochem. Biophys. Res. Commun. 236:178-183(1997).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORMS IIC3; IIC3; BIN1-10-13; BIN1-13 AND  
BIN1-12A).  
RC TISSUE=Fibroblast;  
RX MEDLINE=98058932; PubMed=9395479;  
RA Wechsler-Reya R.J., Sakamuro D., Zhang J., Duhadaway J.,  
RL Prendergast G.C.;  
RL "Structural analysis of the human BIN1 gene. Evidence for tissue-  
RT specific transcriptional regulation and alternate RNA splicing.";  
RL J. Biol. Chem. 272:31453-31458(1997).  
RN [8]  
RP SEQUENCE FROM N.A. (ISOFORM IIC2).  
RA Zhang J., Du W., Wechsler-Reya R.J., Duhadaway J., Sakamuro D.,  
RL Prendergast G.C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP SEQUENCE OF 133-593 FROM N.A.  
RC TISSUE=Brain;  
RX Yu W., Gibbs R.A.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [10]  
RP CHARACTERIZATION.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=98078712; PubMed=9418903;  
RA Wechsler-Reya R.J., Elliott K.J., Prendergast G.C.;  
RL "A role for the putative tumor suppressor Bin1 in muscle cell  
RT differentiation.";  
RL Mol. Cell. Biol. 18:566-575(1998).  
CC -!- FUNCTION: May be involved in regulation of synaptic vesicle  
CC endocytosis. May act as a tumor suppressor and inhibits  
CC malignant cell transformation.  
CC -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).  
CC Binds to SYN1 and DNMI through its SH3 domain, and to clathrin  
CC through a region outside of the SH3 domain. Also binds to alpha-  
CC adaptin. Interacts with the N-terminal transactivation domain of  
CC MYC in a manner requiring the integrity of the conserved MYC box  
CC regions 1 and 2.  
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Isoform Iia is  
CC found in the cytoplasm while isoform BIN1 is nuclear.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=11;  
CC Comment=Additional isoforms seem to exist;  
CC Name=IIA;  
CC IsoId=000499-1; Sequence=Displayed;  
CC Name=IIB;  
CC IsoId=000499-2; Sequence=VSP\_000246, VSP\_000252;  
CC Name=IIC1;  
CC IsoId=000499-3; Sequence=VSP\_000249;  
CC Name=IIC2;  
CC IsoId=000499-4; Sequence=VSP\_000246, VSP\_000249;  
CC Name=IID;  
CC IsoId=000499-5; Sequence=VSP\_000246, VSP\_000247, VSP\_000250;  
CC Name=IIE;  
CC IsoId=000499-6; Sequence=VSP\_000246, VSP\_000253;  
CC Name=IIF;  
CC IsoId=000499-7; Sequence=VSP\_000246, VSP\_000250;  
CC Name=BIN1;  
CC IsoId=000499-8; Sequence=VSP\_000246, VSP\_000247, VSP\_000250;  
CC Name=BIN1-10-13;  
CC IsoId=000499-9; Sequence=VSP\_000246, VSP\_000251;  
CC Name=BIN1-13;  
CC IsoId=000499-10; Sequence=VSP\_000246, VSP\_000247, VSP\_000251;  
CC Name=BIN1-12A;  
CC IsoId=000499-11; Sequence=VSP\_000246, VSP\_000247, VSP\_000253;  
CC -!- TISSUE SPECIFICITY: ISOFORM IIA IS EXPRESSED ONLY IN THE BRAIN  
CC WHERE IT IS CONCENTRATED IN AXON INITIAL SEGMENTS AND NODES OF  
CC RANVIER. ISOFORM BIN1 IS WIDELY EXPRESSED WITH HIGHEST EXPRESSION  
CC IN SKELETAL MUSCLE.  
CC -!- PTM: Phosphorylated by protein kinase C (By similarity).  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF004015; AAC51345.1; -  
CC EMBL; AF070576; AAC28646.1; -  
CC EMBL; U68485; AAC17461.1; -  
CC EMBL; AF001383; AAB61363.1; -  
CC EMBL; AF043898; AAC39210.1; -  
CC EMBL; AF043899; AAC39711.1; -  
CC EMBL; AF043900; AAC39712.1; -  
CC EMBL; AF043901; AAC39713.1; -  
CC EMBL; U87558; AAB63263.1; -  
CC EMBL; AF068914; AAC24126.1; -  
CC EMBL; AF068915; AAC24127.1; -  
CC EMBL; AF068916; AAC24128.1; -  
CC EMBL; AF068917; AAC23750.1; -  
CC EMBL; AF068918; AAC23751.1; -  
CC EMBL; U84004; AAC23440.1; -  
CC EMBL; U83999; AAC23440.1; JOINED.  
CC EMBL; U84001; AAC23440.1; JOINED.  
CC EMBL; U84002; AAC23440.1; JOINED.  
CC EMBL; U84003; AAC23440.1; JOINED.  
CC EMBL; U84004; AAC23441.1; JOINED.  
CC EMBL; U83999; AAC23441.1; JOINED.  
CC EMBL; U84001; AAC23441.1; JOINED.  
CC EMBL; U84002; AAC23441.1; JOINED.  
CC EMBL; U84003; AAC23441.1; JOINED.  
CC PIR: JC5593; JC5593.  
CC TRANSFAC; T03490; -  
CC Genew; HGNC:1052; BIN1.  
CC MIM; 601248; -  
CC GO; GO:0015629; C:actin cytoskeleton; TAS.  
CC GO; GO:0008283; P:cell proliferation; TAS.  
CC GO; GO:0006899; P:nonselective vesicle transport; TAS.  
CC GO; GO:0000074; P:regulation of cell cycle; TAS.  
CC InterPro; IPR003005; Amphiphysin.  
CC InterPro; IPR004148; BAR\_dom.  
CC InterPro; IPR001452; SH3.  
CC Pfam; PF03114; BAR; 1.  
CC Pfam; PF00018; SH3; 1.  
CC PRINTS; PR01251; AMPHIPHYSIN.  
CC PRINTS; PR00452; SH3DOMAIN.  
CC PRODOM; PD000066; SH3; 1.  
CC PROSITE; PS00002; SH3; 1.  
CC KW Alternative splicing; SH3 domain; Coiled coil; Endocytosis;  
CC Anti-oncogene; Differentiation; Phosphorylation.  
CC DOMAIN 15 42 COILED COIL (POTENTIAL).  
FT





```

Db 247 GAPSDGPIRAKTPSPPEERASPIPSPTASPNHTLAPASAPVRPSQTRKGPVPPL 306
QY 292 --VSATEDLADAQ--EDN-----SEIKELLEEEIEKGSSEASSEEDDPL-----PA 338
Db 307 PKVPTKELQENIINFEDFVPEINVTTPSQNEVLEVKETLLDLDPFKPDVTPA 366
QY 339 CNGPAAQPSPTTERAKSQBEVLP-----SST---TSPGALSPSQBSSSA----- 383
Db 367 --GSAATHSP-----MSOTLPWDLWTSTDLVQPASGSGFNDTPQODTSLFTMQTD 417
QY 384 -----TEVLTTRTASE-----GSEQPKKRASIQRTSAPSRPP----- 417
Db 418 QNMAETEQALPTEPCAEPPPTAAAPTAGLDLGLMEEPKEEAAI-----PPTDAGETV 472
QY 418 -----PPRATSPRSGNIPSSPTASGGSGFTSPRASLTGTGTAS----- 457
Db 473 GTEGSGTEAEAEKAAALPAGESGEPGAKIDVESTELAS---SESQAAELEAGAPQEKV 529
QY 458 -PRTSLEVPNPEPEKPVRT--PEAKE-NENIHQNP-----BELCTSPMTLSQVAS-- 507
Db 530 IPSVVIENASNEGHEHQTETTGTTETATEDVAPQCPAGEKQELATEPTPLDSQAATPA 589
QY 508 -----EPGEAKKME---DKEKDK-----LISADSSEGD 534
Db 590 PAGAVDASLSAGDAQAELPPGGLYKVELTHDFAANSDELTLQRGDVVLVWPSDSEADQ 649

RESULT 5
AMPH HUMAN
ID AMPH HUMAN STANDARD; PRT; 695 AA.
AC P49418; O43538;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amphiphysin.
GN AMPH OR AMPHL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX MEDLINE=94357284; PubMed=8076697;
RA David C., Solimena M., de Camilli P.;
RT "Autoimmunity in stiff-man syndrome with breast cancer is targeted to
RT the C-terminal region of human amphiphysin, a protein similar to the
RT yeast proteins, Rvs167 and Rvs161."
RL FEBS Lett. 351:73-79(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=95276740; PubMed=7757077;
RA Yamamoto R., Li X., Winter S., Francke U., Kilimaun M.W.;
RT "Primary structure of human amphiphysin, the dominant autoantigen of
RT paraneoplastic stiff-man syndrome, and mapping of its gene (AMPH) to
RT chromosome 7p13-p14."
RL Hum. Mol. Genet. 4:265-268(1995).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Breast;
RX MEDLINE=98174372; PubMed=9513187;
RA Floyd S.R., Butler M.H., Cremona O., David C., Freyberg Z., Zhang X.,
RA Solimena M., Tokunaga A., Ishizui K., Tsutsui K., De Camilli P.V.;
RT "Expression of amphiphysin I, an autoantigen of paraneoplastic
RT neurological syndromes, in breast cancer."
RL Mol. Med. 4:29-39(1998).
CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in
CC synapses and certain endocrine cell types. May control the
CC properties of the membrane associated cytoskeleton.
CC -!- SUBUNIT: Heterodimer of AMPH and AMPHL (By similarity).
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC SURFACE OF

```

```

CC SYNAPTIC VESICLES.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=128 kDa;
CC IsoId=P49418-1; Sequence=Displayed;
CC Name=2; Synonyms=108 kDa;
CC IsoId=P49418-2; Sequence=VSP_000245;
CC -!- TISSUE SPECIFICITY: Neurons, certain endocrine cell types and
CC spermatocytes.
CC -!- DISEASE: Patients with stiff-man syndrome, a rare disease of the
CC central nervous system characterized by progressive rigidity of
CC the body musculature with superimposed painful spasms, have
CC antibodies against AMPH.
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U07616; AAA21865.1; -.
CC EMBL; X81438; CAA57197.1; -.
CC EMBL; AF034996; AAC02977.1; -.
CC PIR; S62400; S62400.
CC PDB; 1KY7; 12-JUN-02.
CC Genew; HGNC:471; AMPH.
CC MIM; 600418; -.
CC GO; GO:0015629; C:actin cytoskeleton; TAS.
CC GO; GO:0008021; C:synaptic vesicle; TAS.
CC GO; GO:0006897; P:endocytosis; TAS.
CC GO; GO:0007268; P:synaptic transmission; TAS.
CC InterPro; IPR003005; Amphiphysin_1.
CC InterPro; IPR003017; Amphiphysin_1.
CC InterPro; IPR006632; BAR.
CC InterPro; IPR004148; BAR_dom.
CC InterPro; IPR001452; SH3.
CC Pfam; PF01114; BAR; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR01251; AMPHIPHYSIN.
CC PRINTS; PR0452; SH3DOMAIN.
CC ProDom; PD003208; Amphiphysin_1; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00721; BAR; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS00002; SH3; 1.
CC Cytoskeleton; SH3 domain; Coiled coil; Alternative splicing;
CC 3D-structure.
FT DOMAIN 10 83 COILED COIL (POTENTIAL).
FT DOMAIN 144 191 COILED COIL (POTENTIAL).
FT DOMAIN 622 695 SH3.
FT VASSPLIC 425 466 Missing (in isoform 2).
FT /FTId=VSP_000245.
FT SEQUENCE 695 AA; 76256 MW; 7884F75AB75BA357 CRC64;
QY 11 GLFAKQVQKFSRAQEKVLQKGLKAVETKDESPESASNFYQQAEGHKLYKDLNLSA 70
Db 7 GIFAPVQKELNQAQEKVLQKGLKAVETKDESPESASNFYQQAEGHKLYKDLNLSA 66
QY 71 VKVHSSKRVSETLQRIYSSSEWDGHEELKALVWNNDLWEDYERKLADQAVRTMEIYVA 130
Db 67 IKGQEAEMKLTESLHEVTEPDWYGEDVKWGEKCDVLWEDFHQKLVDSLLTLDYLG 126
QY 131 QFSEIKERIAKGRKLVDSAPHHLEAVQNAK-KDEAKTAAEEFNKAQTVFEDLNQE 189
Db 127 QFPDKNIKRAKGRKLVDSAPHHLEAVQNAK-KDEAKTAAEEFNKAQTVFEDLNQE 186

Query Match 28.2%; Score 804.5; DB 1; Length 695;
Best Local Similarity 34.3%; Pred. No. 2.1e-24;
Matches 206; Conservative 104; Mismatches 175; Indels 115; Gaps 18;

```

190 LLEELPILNSRIGCVVTFIOMISNLRDVFYREMSKLNHNLYEVMKLEKOHNSKVFVWK 249  
 187 LOELPILNSRIGCVVTFIOMISNLRDVFYREMSKLNHNLYEVMKLEKOHNSKVFVWK 246  
 250 GLSSSS---RRSLVSPVVRVATVSPPLTS-----PTSPSTLSLKSESESVSATEDLAP 300  
 247 GAPSDSGPLRIAKTSPSPPEPSPPLSPPTASPNHNTLAPASPARPRSPSQT----- 297  
 301 DAAQEDNESEIKELLESEIEKEGSEASSEEDDPLP-ACNGPAQAQPSPTTERAKSQE 358  
 298 ---RKGPVPVPLKVPVTKELQOE---NIISFFDNFVPEISVTTFSQNE---VPEVKKEE 349  
 359 EVL-----PSSTPSPGALSPPSGOPSS-----SATEVLRTRTASSEG---EQP 400  
 350 TLLDLDFDPKPEVTPAGSAGVTHSPMSQTLFWDLWTSTDLV---QPASGSGFNGFTQP 406  
 401 KKRASI-----QRTSAPSPRPPTATASPRPSGNIPSPSTASGSGSFTSP 447  
 407 QDTSFTMTQDOSMICNLAESQAPPTPEPKAEPLAAVTPAVG----- 449  
 448 RASLGTGTASPRTSLEVPSPPEPKPVRTPEAKENENHNQNPDELCTSPTLMTSQVAS 507  
 450 -LDLGMDTRA-----EPPVEAVIIPGADADAAGV-----LVSAEG 486  
 508 EPG---EAKM-----EDKEKNKILISDSSEGQDQLOVMSVPPNNMLTAPPEQEVST 559  
 487 APGEAEAEKATVPAGEVGSLEAKIGTETGAESAQ-----PEAELEATVPOEKVIPS 542

RESULT 6  
 ID AMPH CHICK STANDARD; PRT; 682 AA.  
 AC 250478;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Amphiphysin.  
 GN AMPH  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 ON NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tetra-hybrid; TISSUE=Forebrain;  
 RX MEDLINE=92331604; PubMed=1628617;  
 RA Lichte B., Voh R.W., Meyer H.E., Killmann M.W.;  
 RT "Amphiphysin, a novel protein associated with synaptic vesicles";  
 RL EMBO J. 11:2521-2530(1992).  
 CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in  
 CC synapses and certain endocrine cell types. May control the  
 CC properties of the membrane associated cytoskeleton.  
 CC -!- SUBUNIT: Heterodimer with BIN1. Binds SH3GLB1 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Peripheral membrane protein associated with  
 CC the cytoplasmic surface of synaptic vesicles.  
 CC -!- TISSUE SPECIFICITY: Is abundant in the forebrain and cerebellum.  
 CC It is also found in the adrenal gland, anterior and posterior  
 CC pituitary.  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X60422; CAA42953.1; -;  
 DR PIR; S22700; S22700.

DR InterPro: IPR003005; Amphiphysin.  
 DR InterPro: IPR003017; Amphiphysin\_1.  
 DR InterPro: IPR006832; BAR.  
 DR InterPro: IPR004148; BAR\_dom.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF03114; BAR; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR PRINTS: PR01251; AMPHIPHYSIN.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR ProDom: PD003208; Amphiphysin\_1; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00721; BAR; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR PROSITE: PS00002; SH3; 1.  
 KW Cytoskeleton; SH3 domain; Coiled coil.  
 FT DOMAIN 10 84 COILED COIL (POTENTIAL).  
 FT DOMAIN 144 191 COILED COIL (POTENTIAL).  
 FT DOMAIN 609 682 SH3.  
 SQ SEQUENCE 682 AA; 75205 MW; 61617F494F39EB20 CRC64;  
 Query Match 28.1%; Score 802.5; DB 1; Length 682;  
 Best Local Similarity 33.1%; Pred. No. 2.5e-24;  
 Matches 206; Conservative 118; Mismatches 203; Indels 95; Gaps 19;  
 QY 1 MAEGAGGAGLAFKAVQKXPSRAQEKVLOKLGKAVETKDERFEQASNFYQQAAEGHKL 60  
 DB 1 MADMK---TGIFARNYQKRLNRAQAEKVLQKLGKADKDEQEFEEYVQNFKQAEGRSL 56  
 QY 61 YKDLNKLFLSAVKVHSESKRVSETLQEIYSSWDGHEELKAIYNNNDLLWEDVEEKLAQ 120  
 DB 57 QRELRLYLALIKGMQDASKKLTSLEHYEYDPWYGRDVMKIGKDELMEDPHQKLVG 116  
 QY 121 AVRTMEIYVAQFSSEIKERIAKGRKLVYDSARHLEAVQNAK-KDEAKTAKAESEFNKA 179  
 DB 117 SLLTLDYLGQFPDIKTRIAKRSKLVYDSARHLEALQSSKXKDEGRITKAEEFQXA 176  
 QY 180 QTVFEDLNQELLELPILYNSRIGCVTIFQNTSLNLDVYREMSKLNHNLYEVMKLEK 239  
 DB 177 QKVFEENFDLQELPSLNSRVGYVNTFNKVSLEAKFHKIALLCHLYEVMTKLGD 236  
 QY 240 QHSNKFVFKVGLSSSS---RRSLVSPVVRVATVSPPLTS-----PTSPSTLSKSESE 290  
 DB 237 QHADKFTIQAPSDSGPLAIKATPSPPEVSPPLSPPTASPNHNTLAPASPARPRSPSQT 296  
 QY 291 SVSATEDLAPAAQEDNESEIKELLESEIEKEGSEASSEEDDPLPACNGPAQAQPS-P 349  
 DB 297 ---LRKGPVPVPLKVPVTKELQOE---NIISFFDNFVPEISVTTFSQNEIP 343  
 QY 350 TTERAKS-----QEEVLPSTTSPGALSPPS-----QPSSS----- 382  
 DB 344 ETKVESLDDLPDPFPKPEVSTGVTHSPMSQTLFWDLWTSTSELVQPASSTAFNGFAQD 403  
 QY 383 ATEVLR-----TTASEGSEQPKKRAIORIS-----APSPSPPPRA---TA 423  
 DB 404 TTAFAVQSNENVTETLTAESAELGKLVETPTAAVKEALTAEPDEPTQAAESIEA 463  
 QY 424 SPRSSG--NIPSPSTASGSGS---PTSPRASLGTGTASPRTSLEVPSPPEKPVRT 478  
 DB 464 GDKETTGLAKESVSAAGGAVAVSDSVVAAAGAGAVRTEQEAAGGDKPQGEKDV 523  
 QY 479 EAKENENHNQNPDELCTSPTLMTSQVASPPGAKKEDKDKNKLISAD--SSEGDQDQ 537  
 DB 524 DVSQ-----EKVSIFSVVIEPASNNEGSEEHVIMNESKDAAEAMTQGTDS-E 573  
 QY 538 VSMVPENNNTAPEQEVST 559  
 DB 574 TSQIGSEKAT-----EEIQT 590

RESULT 7  
 ID HOB1 SCHPO STANDARD; PRT; 466 AA.  
 AC 074352;

DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein hbl (Homolog of Binl).  
 DE HOB1 OR SPBC21D10.12.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RX MEDLINE=22457105; PubMed=12569356;  
 RT "hob1", the fission yeast homolog of Binl, is dispensable for  
 RT endocytosis or actin organization, but required for the response to  
 RT starvation or genotoxic stress.;  
 RL Oncogene 22:637-648(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Spouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds S., Squares K., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkheart G., Aert R., Robben J., Grymonprez B.,  
 RA Welltens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehmach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard J., Talla V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomycetes pombe."  
 CC -!- FUNCTION: Has a role in DNA damage signaling as a part of stress  
 CC response processes.  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF275637; AAF86458.1; -;  
 DR EMBL; AL031536; CAA20768.1; -;  
 DR PIR; T11684; T11684.  
 DR GeneDB; SPombe; SPBC21D10.12; -;  
 DR HSSP; P08631; 1BU1.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR dom.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00114; BAR; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.

DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 KW SH3 domain; Coiled coil.  
 FT DOMAIN 7 230  
 FT DOMAIN 294 409  
 FT DOMAIN 407 466  
 FT DOMAIN 31 67  
 FT DOMAIN 177 204  
 FT DOMAIN 466 AA; 51391 MW; 441793950D6C2C7E CRC64;  
 SQ SEQUENCE 466 AA; 51391 MW; 441793950D6C2C7E CRC64;  
 Query Match 8.3%; Score 238; DB 1; Length 466;  
 Best Local Similarity 22.0%; Pred.No.0.0061;  
 Matches 110; Conservative 69; Mismatches 194; Indels 128; Gaps 16;  
 QY 15 KQVOKKESRAQEKVLQKLGKAVETKDERPQSNFYQQQAEGHKLKDKNPLSAVKM 74  
 DB 4 KGFTKALARTPOTLRSENFVGEITKPIYDAGRRKSLTEAKLAEDAKKYTDAINGL 63  
 QY 75 HESSKRVSETLQETIY-----SSEWDGH-----EELKAIWNNDLLWEDYEK 116  
 DB 64 LNHQIGFADACIETKPIISGRASDPSEYEQEAGIEAAEAYKEIVY-----DLQKN 116  
 QY 117 LADQAVRTMEITYAQF-----SEIKERIAKGRKLVVDYSARHHLBAVQNAK-- 163  
 DB 117 LASE-----MDVINTRIVNPTGELLKIVKVDKLLRDKHQLDYDRHRSFKKLQEKDK 172  
 QY 164 --KDEAKTAKAEEFNKAQTVFEDLNQELLELPILYNSRIGCYVTIFQINSLRDVFR 221  
 DB 173 SLKDEKLYEATAFEQSSQSEYETNMLKEELPKLFALAAQSFIAFLFQ-----FY 225  
 QY 222 EMSKLNHLNLYVMSKLEKQHSNFKVVKGLSSSRSLVISPVPVATATVSSPLTSPSPS 281  
 DB 286 MQLNVYVYVLEKMSHCIEIYFD--FNTDILSEYERGDVKDRAELTITTKFAKPTTK 283  
 QY 282 TLSKSESESVSATEDLAPDAAQEDNSEIKELLEEEIEKEGSEASSEEDDPLPACNG 341  
 DB 284 RFGMGPGGKDATAS-----SSSSPSSKREE----- 308  
 QY 342 PAQMGPSPTTTRAKSQEVLPSSTTSPGALSPSGQSSSATEVLRTRTASGSEQPK 401  
 DB 309 -AAAEPSST-----ATDIPPYSTPSVAGASDYS--TPSAGYQTVOTTTTTTEAAAAQY 360  
 QY 402 KRASIQRTSAPSPPPPRATAGPRPSGNIPSSPTAGSGSPTSASLIGTGTSAPRPS 461  
 DB 361 PQARF-----PPPP--VMPQAA-----AAVTFVAAPVAA 389  
 QY 462 LEVS-DNPEPEKPVRTPEAK 481  
 DB 390 AAAAVPVPPAPAPAPAPAAAE 410

## RESULT 8

NFM\_BOVIN STANDARD; PRT; 810 AA.  
 ID NFM\_BOVIN  
 AC 077788;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Neurofilament triplet M protein (160 kDa neurofilament protein)  
 DE (Neurofilament medium polypeptide) (NF-M) (Fragment).  
 GN NEF3 OR NEFM OR NFM.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hill W.D., Zhang L., Balin B.J., Sprinkle T.J., Spicer K.,  
 RA Gearhart D.A.;  
 RT "The bovine neurofilament M subunit has a novel set of KSP repeats  
 normally restricted to NF-H.";

Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate  
 CC filament proteins: L, M, and H which are involved in the  
 CC maintenance of neuronal caliber.  
 CC -!- PTM: There are a number of repeats of the tripeptide K-S-P, NFM is  
 CC phosphorylated on a number of the serines in this motif. It is  
 CC thought that phosphorylation of NFM results in the formation of  
 CC intermediate cross bridges that are important in the maintenance  
 CC of axonal caliber (By similarity).  
 CC -!- PTM: Phosphorylation seems to play a major role in the functioning  
 CC of the larger neurofilament polypeptides (NF-M and NF-H), the  
 CC levels of phosphorylation being altered developmentally and  
 CC coincident with a change in the neurofilament function (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF091342; AAC36357.1; -  
 CC InterPro; IPR001664; IF.  
 CC Pfam; PF00038; Keratin\_1.  
 CC PRINTS; PR01248; TYPE1KERATIN.  
 CC PROSITE; PS00226; IF\_1.  
 CC Intermediate filament; Coiled coil; Neurone; Phosphorylation.  
 KW NON\_TER 1  
 FT DOMAIN <1 296 ROD.  
 FT DOMAIN 297 810 TAIL.  
 FT DOMAIN <1 20 COIL 1A.  
 FT DOMAIN 21 33 LINKER 1.  
 FT DOMAIN 34 132 COIL 1B.  
 FT DOMAIN 133 149 LINKER 12.  
 FT DOMAIN 150 171 COIL 2A.  
 FT DOMAIN 172 175 LINKER 2.  
 FT DOMAIN 176 296 COIL 2B.  
 FT DOMAIN 503 582 8 X 13 AA TANDEM REPEATS.  
 SQ SEQUENCE 810 AA; 90799 MW; B8477D85560AC3F6 CRC64;  
 Query Match 8.1%; Score 232; DB 1; Length 810;  
 Best Local Similarity 21.5%; Pred. No. 0.018;  
 Matches 138; Conservative 98; Mismatches 242; Indels 164; Gaps 25;  
 QY 41 EFPEQASNFYQQOAGHKLKDLNLFSAVKVWHSSKVSSTLQBIYSSEWDGHEELK 100  
 DB 71 EFPEEARLRDTEAIPALRKDIEE-SSLVKV--ELDKV-QSLQDQVAFRLSNHEEV 126  
 QY 101 A-----IWNNDLLWEDYEELAD-----QAVRTMEIYVAQFSE- 134  
 DB 127 ADLLAQIASHITVERKDYLTDTISTALKEIRSQLSHSQDNQHQAEWFKCRYAKLTA 186  
 QY 135 -----IKETIAGKRLVDYDABHHLEAVQAKD-BAKTAEEEFNKATVTF 183  
 DB 187 AEQNKETAIRAKBEIAEYRQL---QSKSTELESVRGTESLERQLSDIEERHNDLSY 243  
 QY 184 EDLNQELLEELPILYNSIGCYVTIFQINSLNLDVFRMSKLNHLYEVMWKLKQHSN 243  
 DB 244 QDTIQLENELRGT-KWMAHRLREYQDLLNVKALDIEIAAYR---KLLEGEETRFST 298  
 QY 244 KVVVKV-GSSSSRRSLVSPVTRATVSSP-----L 274  
 DB 299 FAGSITGLYTHROPSIAISSKIQKTVKAPKLQVQHKFVEIEETKVEDEKSEMEAL 358  
 QY 275 TSPTSPTLSLKSE-----SESVATEDL-----APDAQGEDNSIKEL 314  
 DB 359 TAITELAVSVKVEEAEKEKEAEVEEVAAKKSPVKATAPTELKEEKEEKEEG 418  
 QY 315 LEEIEKEGEASSEEDDPLPACNGPAQAPQSPFTTTRAKSQEVLSPSTTPPGGALS 374

## RESULT 9

NFM\_HUMAN  
 ID NFM\_HUMAN STANDARD; PRT; 915 AA.  
 AC P07197;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Neurofilament triplet M protein (160 kDa neurofilament protein)  
 DE (Neurofilament medium polypeptide) (NF-M) (Neurofilament 3).  
 GN NEF3 OR NEFM OR NFM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8725853; PubMed=3608989;  
 RA Myers M.W., Lazzarini R.A., Lee V.M.-Y., Schlaepfer W.W., Nelson D.L.;  
 RT "The human mid-size neurofilament subunit: a repeated protein  
 RT sequence and the relationship of its gene to the intermediate  
 RT filament gene family."  
 RL EMBO J. 6:1617-1626(1987).  
 RN [2]  
 RP SEQUENCE OF ONE OF THE 13 RESIDUE REPEATS.  
 RX MEDLINE=88158120; PubMed=2450354;  
 RA Lee V.M.-Y., Otvos L. Jr., Carden M.J., Hollosi M., Dietzschold B.,  
 RA Lazzarini R.A.;  
 RT "Identification of the major multiphosphorylation site in mammalian  
 RT neurofilaments."  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1998-2002(1988).  
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate  
 CC filament proteins: L, M, and H which are involved in the  
 CC maintenance of neuronal caliber.  
 CC -!- PTM: There are a number of repeats of the tripeptide K-S-P, NFM is  
 CC phosphorylated on a number of the serines in this motif. It is  
 CC thought that phosphorylation of NFM results in the formation of  
 CC intermediate cross bridges that are important in the maintenance  
 CC of axonal caliber.  
 CC -!- PTM: Phosphorylation seems to play a major role in the functioning  
 CC of the larger neurofilament polypeptides (NF-M and NF-H), the  
 CC levels of phosphorylation being altered developmentally and  
 CC coincident with a change in the neurofilament function.  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Y00067; CAA68276.1; -.

Db 419 QEEEEEAAKQDAEE-----GGSEKESSEKEGEQEE-----GETAEGESE 466  
 QY 375 PSQGP-----SSSATEVLTTRTASEGSPQPKGRASIQRTSAPPPPPPPRATASPRPS 428  
 Db 467 AAABAKKEKMEKAEVAPKEELAAEKAEKPEKAPVAKSPPTTKSPPTAKSPEAKSPE 526  
 QY 429 SGNTPSPSTASGGSP--SPRASLTGTGTASPTSLVSPNPPPEPKPVRTPEAKENENI 486  
 Db 527 A-----KSPTAK---SPTAKSPVAK-----SPTAKSPEAKSPEAKSPTAKSP 568  
 QY 487 HNQNPBELCTSPTLMTSQV-----ASEPGRK-KMEDKEKDKNLISADSSEGDQLOVS 539  
 Db 569 ---SPAASKPAPKSPVEEVKPAEAGAEKGEQKEVEEKKEAKESPKKEKAEKKEKPK 625  
 QY 540 MYPENNNLTAP-----EPQEVSTSENPO 563  
 Db 626 DVPEKKKAEKSPVKAESPVKEVPKPVKVSPEKAEKEEKPKQ 667

DR A27864; A27864.  
DR Genew; HGNC:7734; NEF3.  
DR MIM; 162250; -  
DR GO; GO:0005883; C:neurofilament; TAS.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.  
DR InterPro; IPR006821; Filament\_head.  
DR InterPro; IPR001684; IF.  
DR InterPro; IPR002957; Keratin\_I.  
DR Pfam; PF00038; filament; 1.  
DR Pfam; PF04732; filament head; 1.  
DR PRINTS; PR01248; TYPE1KERATIN.  
DR PROSITE; PS00226; IF; 1.  
DR PROSITE; PS00226; IF; 1.  
KW Intermediate filament; Coiled coil; Neurone; Phosphorylation;  
KW Glycoprotein.  
FT INIT\_MET 0 0  
FT DOMAIN 1 103 HEAD.  
FT DOMAIN 104 411 ROD.  
FT DOMAIN 412 915 TAIL.  
FT DOMAIN 104 135 COIL 1A.  
FT DOMAIN 136 148 COIL 1B.  
FT DOMAIN 149 247 LINKER 1.  
FT DOMAIN 248 284 LINKER 2.  
FT DOMAIN 285 286 COIL 2A.  
FT DOMAIN 287 290 LINKER 2.  
FT DOMAIN 291 411 COIL 2B.  
FT DOMAIN 613 690 6 X 13 AA TANDEM REPEATS.  
FT CARBOHYD 46 46 O-LINKED (GLCNAC) (BY SIMILARITY).  
FT CARBOHYD 430 430 O-LINKED (GLCNAC) (BY SIMILARITY).  
FT SEQUENCE 915 AA; 102317 MW; 5F5D3DF34C9D9B50 CRC64;  
Query Match 8.1%; Score 230.5; DB 1; Length 915;  
Best Local Similarity 22.4%; Pred. No. 0.023;  
Matches 131; Conservative 96; Mismatches 264; Indels 93; Gaps 24;  
QY 41 ERFEQASNFYQQAGHKLKYLKXNPLSAVKVWHSSKRYSETLQIYSEWDGHEELK 100  
DB 186 ERFEARLRDTEAARLRKDIEEASLVKVV--ELDKV-QSLQDEVAFLSNHEEV 241  
QY 101 AIJWNNDLWE-----DYEKLAQAVRTMEIYVAQFS-----EIKRIAKGR 144  
DB 242 A-----DLIAQIQAASHITVERKDYLTIDTALKEIRSQLSHSDQNHQAQEFKRYA 296  
QY 145 KLVDYDSARHLEAVQNAKDEAKTAAEE---EFNKAQTVFDINQELLEPLILYNS 200  
DB 297 KLT--EAAQNKKEATRSABETAEYRQIQSKSIELESVRGKESLREQ--LSIDIEERHNH 353  
QY 201 RIGCVYTIQINSLNRDVFYREMSKLNHNLYEVMSKL--EKSHNKNVFWKGLSSRRSL 259  
DB 354 DISSYDITQLEN-----ELRCKVEMARHLREYQDLNVKQALDIIAYRKL 403  
QY 260 VISPVPVATVSSPTSP-----TSPTSLSKSESVSA-----TBDLAPDAQGE 306  
DB 404 LRGESTRFTFAGSTGLYTHRPITITSSIKQTKVAPKLKVQHKFVFEIETKVED 463  
QY 307 DNSIEKELL-----EIEKEGEASSE---EDDPLPACNGPAQA--QPSPTTER 353  
DB 464 EKSEMEALTAITELASMKKEEKAEEKEEPEAEAEVAAKSPVATAPVKEEE 523  
QY 354 AKSQEVLPSSVTPFGGALSPGQPSGSSATEVVL-----RRTTASEGE--QPKK 402  
DB 524 GEKEEGEGEBEEDEGAKSQAEEGSEKESGEQEGEGBEGEGBEGEAEAEAE 583  
QY 403 RASIORSTAPPSPPPPPRATA---SPRSSGNIPSSPTASGGSPPT--SPRSLGTCGTAS 457  
DB 584 EKKVEKSEVATKEELVADAKVEKPAKSPVPSVPEEKSPVPSVPEEKSPV- 642  
QY 458 PRISLEVSNPFPPEKPVRTPEAKENIHNQNPEELCTSP-----LMTSQVASEPGEAK 513  
DB 643 PKSPVEEKSPVPSKSPV---EEKKSPVSKSPVPEEKAKSPVPSVPEEKAKSVEYKGE 699  
QY 514 KMEDEKONKLLISAQSSGQDQLQVSMVPENNNTAPEPQEVUS 557  
DB 700 QKEEEKVEKAPKEKVEKKEKPKDVPKPKKAEKSPVKEEAVA 743



FT DOMAIN 243 259 LINKER 12.  
FT DOMAIN 260 281 COIL 2A.  
FT DOMAIN 262 285 LINKER 2.  
FT DOMAIN 286 406 COIL 2B.  
FT CARBOHYD 46 46 O-LINKED (GLCNAC) (BY SIMILARITY).  
FT CARBOHYD 46 46 O-LINKED (GLCNAC) (BY SIMILARITY).  
FT CARBOHYD 426 426 O-LINKED (GLCNAC) (BY SIMILARITY).  
FT CONFLICT 546 546 G -> R (IN REF. 2).  
SQ SEQUENCE 857 AA; 95704 MW; 452E0FC6AC64778B CRC64;

Query Match 8.0%; Score 228.5; DB 1; Length 857;  
Best Local Similarity 21.2%; Pred. No. 0.026;  
Matches 149; Conservative 91; Mismatches 276; Indels 187; Gaps 27;

QY 18 QKFSRAQ-----EKVLQKLGKAVETKD-----ERFQSA-----47  
DB 132 QKHAGRAQLGDAYQELFELGALAEQVSHKQAQIQDSEHIEDIOIRLERFDEARLSD 191  
QY 48 -----SNFYQQQAEGHLYKDLKNFLSAVKVNHSSKXVSEPLQIYSEWD 94  
DB 132 ETEATIALRKEMEEASLMRAELDKVQSLQDEVAFLRGHNE--EVAELLAQLQASH-- 247  
QY 95 GHEELKAIWNNDLLMEDYEKLAD-----QAVRTMEIYVAQFSE-----134  
DB 248 -----ATVERKDYKTLDTLTALKEIRAQLQECSDHNMHQAEEWFKRYAKLTEAAEQNK 301  
QY 135 -----IKERIAKGRKLVDSARHLEAVQNAKQD-EAKTAKAEFEFNKAQTVFED-----185  
DB 302 EAIRSAKEEIAEYRQL-----QSKSIELESVRGKTESLERQLSDIEBRHNDLTYQDTTH 358  
QY 186 -LNOEL-----LEELPILNSR-----IGCVTLIFONISNLROVF-----219  
DB 359 QLENELRTGKWMARHLREYDLDLVNKMALDIEIAVYKLLGEEETFPFSGSITGPIF 418  
QY 220 -YREMS-----KLNHLYEVM-----SKLEKQHSNK-----VFVVKGLSS 253  
DB 419 THRQPSVTIASTKIQTKEPKLVQHKFVEEIIETBEKVEDEKSEMEDALSALAEEMAA 478  
QY 254 SSRSLVISPVRVATVSSPLTSPSTSLSKSESESUSATDLA--PDAQ--CEDNS 309  
DB 479 KAQEEQEEKAE 538  
QY 310 EIKELLEEEETKSGSEASSEED-----DPLPACNGPAQAQFSPPTTERAKSSEV 360  
DB 539 EIEEKEGEAE 597  
QY 361 L-----PSTTSPGGLSPSCQSSSATEVVLTRTASESEQPKRASIQTSP- 412  
DB 598 AAEVKGQKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 653  
QY 413 ----PSRPPPPPRATASP-RPSSGNIPSSPTASGGGSPSPRASLGVTGASPTSLSEVSN 467  
DB 654 KVRSPKPTTPEKVVSPKSPKPTPEKP--ASPEKPAPEKPTPEKPAPEKPS 711  
QY 468 PEPKPVPTPEAKENINHNQNPDELCTSPMLTSQV-----ASEPGEAKKMD 517  
DB 712 PEKESPLKDEKAVVEESI-----TVTKVTKVTAEEVSEKAEKEDIANVEGEKKD 764  
QY 518 KEKDKNLISADSSGDOLOVSMVPPNNNLTAPEPQEEVSTSE 560  
DB 765 EAKEKEAEKEKGVVTVNGLVSPVDEKGEKVVTKAEKITSE 807

RESULT 11  
NCR2 HUMAN  
ID NCR2 HUMAN STANDARD; PRT: 2517 AA.  
AC Q9V618; Q00613; Q15416; Q13354; Q9V5U0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Nuclear receptor co-repressor 2 (N-CoR2) (silencing mediator of  
DE retinoic acid and thyroid hormone receptor) (SMRT) (Thyroid-,  
DE retinoic-acid-receptor-associated co-repressor) (T3 receptor-  
DE associating factor) (TRAC) (CTG repeat protein 26).

GN NCR2 OR CTG26.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM SMRT).  
RC TISSUE=Pituitary;  
RX MEDLINE=99178941; PubMed=10077563;  
RA Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;  
RT "Unique forms of human and mouse nuclear receptor corepressor SMRT.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:2639-2644(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM SMRT).  
RC TISSUE=Cervical adenocarcinoma;  
RX MEDLINE=9919215; PubMed=10097068;  
RA Park E.J., Schroen D.J., Yang M., Li H., Li L., Chen J.D.;  
RT "SMRT, a silencing mediator for retinoid and thyroid hormone  
RT receptor-extended isoform that is more related to the nuclear  
RT receptor corepressor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524(1999).  
RN [3]  
RP SEQUENCE OF 1023-2517 FROM N.A.  
RC TISSUE=Cervical adenocarcinoma;  
RX MEDLINE=96008552; PubMed=7566127;  
RA Chen J.D., Evans R.M.;  
RT "A transcriptional co-repressor that interacts with nuclear hormone  
RT receptors.";  
RL Nature 377:454-457(1995).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM TRAC-1).  
RC TISSUE=Fetal liver;  
RX MEDLINE=96408715; PubMed=8813722;  
RA Sande S., Privalsky M.I.;  
RT "Identification of TRACs (T3 receptor-associating cofactors), a family  
RT of cofactors that associate with, and modulate the activity of,  
RT nuclear hormone receptors.";  
RL Mol. Endocrinol. 10:813-825(1996).  
RN [5]  
RP SEQUENCE OF 428-613 FROM N.A.  
RC TISSUE=Brain cortex;  
RX MEDLINE=97369492; PubMed=9225980;  
RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kidwai A.S.,  
RA Breachel T.S., Stine O.C., Callahan C., McInnis M.G., Ross C.A.;  
RT "cDNAs with long 5' UTRs that contain repeats from human brain.";  
RL Hum. Genet. 100:114-122(1997).  
RN [6]  
RP INTERACTION WITH MINT.  
RX MEDLINE=21231190; PubMed=11331609;  
RA Shi Y., Downes M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,  
RA Hon M., Evans R.M.;  
RT "Sharp, an inducible cofactor that integrates nuclear receptor  
RT repression and activation.";  
RL Genes Dev. 15:1140-1151(2001).  
RN [7]  
RP INTERACTION WITH HDAC10.  
RX MEDLINE=21839031; PubMed=11739383;  
RA Fischer D.D., Cai R., Bhatia U., Asselbergs F.A.M., Song C., Terry R.,  
RA Trogiani N., Widmer R., Atadja P., Cohen D.;  
RT "Isolation and characterization of a novel class II histone  
RT deacetylase, HDAC10.";  
RL J. Biol. Chem. 277:6656-6666(2002).  
CC -|- FUNCTION: Mediates the transcriptional repression activity of some  
CC nuclear receptors by promoting chromatin condensation, thus  
CC preventing access of the basal transcription.  
CC -|- SUBUNIT: Interacts with HDAC7 (By similarity). Forms a large  
CC corepressor complex that contains SIN3A/B and histone deacetylases  
CC HDAC1 and HDAC2. This complex associates with the thyroid (TR) and  
CC the retinoid acid receptors (RAR) in the absence of ligand, and  
CC may stabilize their interaction with TR/RAR. The SMRT isoform  
CC interacts with HDAC10. Interacts with MINT.  
CC -|- SUBCELLULAR LOCATION: Nuclear.  
CC -|- ALTERNATIVE PRODUCTS:



GN NEF3 OR NEFM OR NFM.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=97055255; PubMed=8999542;  
 RT Vitadello M., Vettore S., Lamar E., Chien K.R., Gorza L.;  
 RT "Neurofilament M mRNA is expressed in conduction system myocytes of  
 RT the developing and adult rabbit heart.";  
 RL J. Mol. Cell. Cardiol. 28:1833-1844(1996).  
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate  
 CC filament proteins: L, M, and H which are involved in the  
 CC maintenance of neuronal caliber.  
 CC -!- PTM: There are a number of repeats of the tripeptide K-S-P. NFM is  
 CC phosphorylated on a number of the serines in this motif. It is  
 CC thought that phosphorylation of NFM results in the formation of  
 CC interfilament cross bridges that are important in the maintenance  
 CC of axonal caliber.  
 CC -!- PTM: Phosphorylation seems to play a major role in the functioning  
 CC of the larger neurofilament polypeptides (NF-M and NF-H). The  
 CC levels of phosphorylation being altered developmentally and  
 CC coincident with a change in the neurofilament function.  
 CC -!- SIMILARITY: Belongs to the intermediate filament family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; Z47378; CAA87454.1; --  
 DR PIR; S55395; S55395.  
 DR InterPro; IPR001664; IF.  
 DR Pfam; PF00038; filament; 1.  
 DR PROSITE; PS00226; IF; 1.  
 KW Intermediate filament; Coiled coil; Neurone; Phosphorylation;  
 KW Glycoprotein.  
 FT NON TER 1 1  
 FT DOMAIN <1 197 ROD.  
 FT DOMAIN 198 643 TAIL.  
 FT DOMAIN <1 33 COIL 1B.  
 FT DOMAIN 34 50 LINKER 12.  
 FT DOMAIN 51 72 COIL 2A.  
 FT DOMAIN 73 76 LINKER 2.  
 FT DOMAIN 77 197 COIL 2B.  
 FT CARBOHYD 217 217 O-LINKED (GLCNAC) (BY SIMILARITY).  
 SQ SEQUENCE 644 AA; 72450 MW; 030FDA622898678 CRC64;  
 Query Match 7.4%; Score 212; DB 1; Length 644;  
 Best Local Similarity 20.9%; Pred. No. 0.083;  
 Matches 121; Conservative 84; Mismatches 244; Indels 130; Gaps 22;  
 QY 62 KDLKNFLSAVKVMESSKR-----VSETLQRIYS-----SEWDGHEELKAIWVNDLL 109  
 DB 25 EVADLLAQIQASHITVERKDYLTDTISSALKIRSOLECHSDQNHHQ-----AEW----- 76  
 QY 110 WEDYEKLDAQVTRMEIYVAQFSEIKERIAKGRKLVYDVSARHLEAVONAKKDEAKT 169  
 DB 77 FKCRYAKLTAEEQNKRAISAKKEEIAEYRQLOSKSIELESVAMHKSLEHRVSD---- 132  
 QY 170 AKAEENFKATQVEDLNQELLEPLLYNSRIGCVVTFPQINISNRDVFYREMSKLNHN 229  
 DB 133 --IBERNHNDUSSYQDTIQLENEL-----RGTKWMAHRLRE--YQDL--LN-- 174  
 QY 230 LYEVMKLEKQHSNKNFVWKLSSSSRRSLVISPPTVATVSSPLTP-----TSPS-TLS 284  
 DB 175 -----VKMALDIEIAVRYKLLGEETRFSTFGSITGPLYTHRQPSVTIS 219

QY 285 LKSESESVSA-----TEDLPAAQGEDNSSEIKELL-----EEEEIEK 322  
 DB 220 SKIQTKEAPKLVQKHKFVEEIEETKVEDESEMEDALTATAELAVSVKKEEKEEA 279  
 QY 323 EGSEASSSEEDDPLPACNGPAQA-QFSPTTTERAKSQEEVLPSSTTSPGCGALSPSQPSS 381  
 DB 280 EGKEEQEAEAEVAAAKKSPKATTPRIKEEKEEKEEKEEEDDEGKVDQAERG 339  
 QY 382 SATWVLT-----RTASEG-----SEQPKRASIOQTAPPSPRPPTATASPRPS 428  
 DB 340 SEKEGSKNGEGEGETEAEAGVEEAEAKKTEKEEVEAAKEEPTVEATKVGKPEKA 399  
 QY 429 SGNIPSSPTASGGGSPTSPPASLIGTGTASPTSL-----VSPNPE-----PPEK 473  
 DB 400 KSPVFKSPV-----EEVKPAEATAGKGEKEEVEEKKKAKEKSPKEKVKKEEK 454  
 QY 474 PVRTPEAKENHNQNPBELC--TSPTLTWTSOVAEPGEAKMKDEKDKNLISADS-S 530  
 DB 455 PDVPKPKKAEKSPKKEAAEAATITPTKVGLEKTKGKPLQOEKEKEKAGEGSGSE 514  
 QY 531 EGQDQ-----LOVSMVPENNLTAPQEPREVSTSE 560  
 DB 515 EGDQSGSKRAKEDIANGEGEKEEKEEPEETKKGSGRE 553  
 RESULT 13  
 HOB3 SCHPO STANDARD; PRT; 264 AA.  
 ID HOB3 SCHPO  
 AC Q9TUM7; P78850;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein hcb3 (Homolog of Bin3).  
 GN HOB3 OR SPBC725.09C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RX MEDLINE=21293007; PubMed=11274158;  
 RA Routhier E.L., Burn T.C., Abbaszade I., Summers M., Albright C.F.,  
 RA Prendergast G.C.;  
 RT "Human BIN3 complements the F-actin localization defects caused by  
 RT loss of Hob3p, the fission yeast homolog of Rvel61p.";  
 RL J. Biol. Chem. 276:21670-21677(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Haunlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkhardt G., Aert R., Robben J., Grymonprez B.,  
 RA Weijens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas R., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

```

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs.";
RL DNA Res. 4:363-369(1997).
CC -!- FUNCTION: Involved in cytokinesis and septation where it has a
CC role in the localization of F-actin.
CC -!- SIMILARITY: Contains 1 BAR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF275638; AAF86459.1; -
CC EMBL; AL034352; CAA22181.1; -
CC FIR; T40661; T40661.
CC GenDB SPombe; SPSC725.09c; -.
CC InterPro; IPR003005; Amphiphysin.
CC InterPro; IPR006632; BAR.
CC Pfam; PF03114; BAR; 1.
CC PRINTS; PRO1251; AMPHIPHYSIN.
CC SMART; SM00721; BAR; 1.
CC
CC Separation; Cytoskeleton; Coiled coil.
FT DOMAIN 25 65
FT COILED COIL (POTENTIAL).
FT DOMAIN 165 187
FT COILED COIL (POTENTIAL).
FT CONFLICT 93 93
FT CONFLICT 96 96
FT SEQUENCE 264 AA; 30094 MW; 845388592305AD4D CRC64;
SQ
Query Match 7.4%; Score 210; DB 1; Length 264;
Best Local Similarity 26.6%; Pred. No. 0.04;
Matches 63; Conservative 44; Mismatches 110; Indels 20; Gaps 6;
QY 18 QKFSRAQKVLQKLGKAVETKDRFRQASNFYQQQAEGHKLYKLNFLSAVKYMHPS 77
Db 7 KKANVRACTSYMKTGHTVDEFEFETERRRYTMSAAKQKQKAGYLDALRAMTAS 66
QY 78 SKRVSETLQETYSSEWDGHEELKALVNNDLWEDYEE---KLADQAVTMEI-----YV 129
Db 67 QTRIANITDAFYGDA--GSKDGVAYTRQ--VVEDLDATVTKELDGGFFRTVLDPISRFC 122
QY 130 AQFSIEKRIAKRGKLVYDSARHLE-AVQNAKDEAKTAKEEFNKAQTVFEDLNQ 188
Db 123 SYFPDINAATKRNHKLHDHMAKQKLVKDKPSNDTTKLPRTKEAAMAKEVYETLNN 182
QY 189 ELLELPILYNSRGCVYTIQNTSLNLDVYREMSKLNHLYEVMSKLEKQHSNKV 245
Db 183 QLVSELPOLIALRVPLDPSFEALVKTLQRFREG-----YERAAQVQQVQVDSNV 232

```

## RESULT 14

```

ID ASPI_HUMAN
AC Q96KQ4; O94870; STANDARD; PRT; 1090 AA.
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Apoptosis stimulating of p53 protein 1 (Protein phosphatase 1
DE regulatory subunit 13B).
GN PPIR135 OR ASFP1 OR KIAA0771.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, DISEASE, AND
RP INTERACTION WITH TP53.
RX MEDLINE=21541920; PubMed=11694014;
RA Samuels-Lev V., O'Connor D.J., Bergamaschi D., Trigliente G.,
RA Hsieh J.-K., Zhong S., Campargue I., Naumovski L., Crook T., Lu X.;
RT "ASP proteins specifically stimulate the apoptotic function of p53.";
RL Mol. Cell 8:781-794(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12508121;
RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
RA Catolico L., Levy M., Barbe V., De Bernardis V., Ureta-Vidal A.,
RA Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,
RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
RA Brulets T., Jaillon O., Friedlander L., Samson G., Brottier P.,
RA Cure S., Segurens B., Antier F., Samain S., Crespeau H., Abbasi N.,
RA Alich N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedman C.,
RA Gouyvenoux M., James R., Madan A., Mairey-Estrada B., Mangenot S.,
RA Martins N., Menard M., Oztas S., Ratcliffe A., Shafer T., Trask B.,
RA Vacherie B., Bellemere C., Belser C., Besnard-Gonnest M.,
RA Bartol-Mavel D., Boutard M., Briez-Silla S., Combette S.,
RA Dutoise-Laurent V., Ferron C., Lechaplais C., Louesse C., Muselet D.,
RA Magdelenat G., Pateau E., Petit E., Sirvain-Trukiewicz P., Trybou A.,
RA Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,
RA Dumont C., Guerin T., Hafray S., Hammadi R., Munga J., Pellouin V.,
RA Robert D., Wundlerle E., Gauguier G., Roy A., Sainte-Marthe L.,
RA Verdier J., Verdier-Discaia C., Hillier L.W., Fulton L., McPherson J.,
RA Matsuda F., Wilson R., Scarpelli C., Gyapay G., Wincker P., Saurin W.,
RA Quetier F., Waterston R., Hood L., Weissbach J.;
RT "The DNA sequence and analysis of human chromosome 14.";
RL Nature 421:601-607(2003).
RN [3]
RP SEQUENCE OF 143-1090 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
CC -!- FUNCTION: Regulator that plays a central role in regulation of
CC apoptosis via its interaction with p53/TP53. Regulates TP53 by
CC enhancing the DNA binding and transactivation function of TP53 on
CC the promoters of proapoptotic genes in vivo.
CC -!- SUBUNIT: Interacts with TP53.
CC -!- SUBCELLULAR LOCATION: Predominantly cytoplasmic. Some fraction is
CC nuclear.
CC -!- DOMAIN: The ankyrin repeats and the SH3 domain are required for a
CC specific interactions with TP53.
CC -!- DISEASE: Defects in PPIR135 may be a cause of breast cancers. It
CC is overexpressed in many patients suffering from breast carcinomas
CC and expressing a wild-type TP53 protein.
CC -!- MISCELLANEOUS: In contrast to its official gene name, it is not a
CC regulatory subunit of protein phosphatase 1. This name was given
CC due to its similarity with a protein that binds to protein
CC phosphatase 1.
CC -!- SIMILARITY: Belongs to the ASPP family.
CC -!- SIMILARITY: Contains 2 ANK repeats.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```
CC EMBL; AJ318887; CAC83011.2; -.
DR EMBL; AL049840; -. NOT_ANNOTATED_CDS.
DR EMBL; AB018314; BAA34431.1; -.
DR Genew; HGNC:14950; PPRI13B.
DR MIM; 606455; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00023; ank; 2.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00248; ANK; 2.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50002; SH3_1.
KW Apoptosis; Nuclear protein; Repeat; ANK repeat; SH3 domain.
FT REPEAT 920 952
FT REPEAT 953 985
FT DOMAIN 1019 1081
FT DOMAIN 132 292
FT DOMAIN 428 860
FT DOMAIN 1090 AA; 119585 MW; 92CF98EDDDDC89A4 CRC64;
SQ SEQUENCE 1090 AA; 119585 MW; 92CF98EDDDDC89A4 CRC64;

Query Match 7.3%; Score 208.5; DB 1; Length 1090;
Best Local Similarity 22.5%; Pred. No. 0.19;
Matches 137; Conservative 85; Mismatches 224; Indels 163; Gaps 31;

QY 12 LFAQGVKTSRAQEKV--LQKLGKAVETKDEPEQSASNFVQQQAGHGLYKDLKFLS 69
DB 126 LTLSELQMAARQQQLNQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 69
QY 70 AVKMVHSSKRVSTLQEI--YSSEWDGH-----EELKAIWNNDLWEDYEKLDAQVR 123
DB 182 RVZAEQNKLRKIRAMRQVDYKIMGNLSAEIERFSAM-----FQEK--KQEVQ 229
QY 124 TWEIYVAFSEIKERIAKGRKLVVDYSARHLEAVQNAKDEAKTAAEEBENKAQTVF 183
DB 230 TAILRVDQLSQQLDEL-KKG-KLNGFQS-----YNGKLTPAAVELKRLYQLQ-IR 278
QY 184 EDLNOE---LLEELPILYNIRIGCYVTFQINLNRDVFYREMSKLNHLYEVMKLEKQ 240
DB 279 NQLNQEQNSKLOQQQLKLNKRNEMVAMDKRISLERLYGKQILNR----- 326
QY 241 HSNKVFVVKGLS-----SSRSLVSPVPRVATVSS--PLTSPSPSTLSKSES--- 289
DB 327 -----VNGTSSPQSPPLSTSGRVAAGVPIQVPSAGSPVLGDPKQPSLSIASNAAG 379
QY 290 ESVSATEDLAPDAAQGDENSEIKEL---LEEEIEKEGSEASSSEEDDPLP-ACNGPAQ 344
DB 380 RSKSANDGNWPTLKQ--NSSSVKPVQVAGADWKDPSVEGSKQTVSSQPVFPFSALGPT 438
QY 345 AQPSPPTTERAKSQEV-----LFSS--TTPSPGGLSPSGQPSSSSATEVVLRTASEG 396
DB 439 ---KPGIEICKVPPPIPGVGKQLPSPYGVFSP-----TLPGFGSTSLSE----- 480
QY 397 SEQPKKRASTQRTSAP-PSRPPP---PRATASPRPSG-----NIPSSPTASGGSGPT 445
DB 481 ---RRKEGSLPRPSAGLPSRQRTLLPATGTPQPGSSQIQIRISVPPSPPTYPAGPPA 537
QY 446 SPRASLTGTASPTSLVSPNP-----EPPEKPVRTPEAKENENINQNPEEL----- 494
DB 538 FP-----AGDSKPELPTVAIRPFLADKGRSQSPFKGPTVNSSIYMYLQQATPPKN 592
QY 495 -----CTSP-----TLMTSQVASEFCEAKCKEDKEKQ 521
DB 593 YQPAASHALNKSVAKVGVKPLPSGSTSPSLPFLHGLSLSTGTPQPPSPSSTEKEPQD 652
QY 522 NKLISADSS 530
DB 653 GPAAPADGS 661
```

```
RESULT 15
SYJL_BOVIN
ID SYJL_BOVIN STANDARD; PRT; 1324 AA.
AC O18964;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Synaptotagmin 1 (EC 3.1.3.36) (Synaptic inositol-1,4,5-trisphosphate 5-phosphatase 1) (p150) (Fragment).
GN SYNJ1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 321-339 AND 454-469.
TI TISSUE=Brain;
TX MEDLINE=97342621; PubMed=9199318;
RA Sakisaka T., Itoh T., Miura K., Takenawa T.;
RT "Phosphatidylinositol 4,5-bisphosphate phosphatase regulates the rearrangement of actin filaments.";
RL Mol. Cell. Biol. 17:3841-3849(1997).
CC -!- FUNCTION: Hydrolyzes PI(2) bound to actin regulatory proteins resulting in the rearrangement of actin filaments downstream of tyrosine kinase and ASH/GRB2.
CC -!- CATALYTIC ACTIVITY: 1-phosphatidy-1D-myo-inositol 4,5-bisphosphate + H(2)O = 1-phosphatidy-1D-myo-inositol 4-phosphate + phosphate.
CC -!- SUBUNIT: Binds to AMPH and ASH/GRB2.
CC -!- SUBCELLULAR LOCATION: Predominantly concentrated in the perinuclear areas.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with highest levels in brain.
CC -!- DOMAIN: The C-terminal proline-rich region mediates binding to a variety of SH3 domain-containing proteins including AMPH and ASH/GRB2.
CC -!- SIMILARITY: In the central section; belongs to the inositol-1,4,5-trisphosphate 5-phosphatase family.
CC -!- SIMILARITY: Contains 1 SAC domain.
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -!- CAUTION: This is a conceptual translation; a number of potential framehifts were corrected starting in position 1213 so as to extend the similarity with the orthologs.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; D85682; BAA31652.1; ALT_FRAME.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000300; IPFC.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR002013; SYJA_N.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF02383; SYJA_N; 1.
DR SMART; SM0128; IPFC; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS00275; SAC; 1.
KW Hydrolase; Endocytosis; RNA-binding; Multigene family.
FT DOMAIN 119 442
FT DOMAIN 475 859
FT DOMAIN 902 971
FT DOMAIN 860 1212
FT DOMAIN 1033 1036
FT DOMAIN 1108 1113
FT DOMAIN 1126 1129
```

```

FT CONFLICT 335 335 Y -> YY (IN REF. 1; AA SEQUENCE).
FT NON TER 1324 1324
SQ SEQUENCE 1324 AA; 146582 MW; EDDC2DD9D6E3191C CRC64;

Query Match
Best Local Similarity 20.9%; Score 204.5; DB 1; Length 1324;
Matches 116; Conservative 79; Mismatches 220; Indels 139; Gaps 23;

QY 87 EYSSWDGHEELKAIWNNDLLWEDYEKLADQAVRTWETIYVAQF-----132
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
787 DLFSDDYDTSEKCRTPAWTRVLRKRWPF--DRSAEDLDLLNASFQDESKILYTWTPGT 845
QY 133 -----SEIKERIAKGRKLVYDSARHLEAVQNAKD-----EAKTAKA 172
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
846 LHYGRRELKTSOHPVVALIDIIDIFEVEAEERQNIYKEVIAVQGPDPGTVLVSIKSLP 905
QY 173 EEFNKAAQTVFEDLNQELLEELPILYNSRIGCYVTIFQNISNLRDV-----F 219
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
906 ENNF-----FNDALIDEL-----LQFTNFGEVILIRFVEDKMWVTF 942
QY 220 YREMSKLNHLYEVMKLEKQHSNKVFVKGLSRSSRLVISPPVRTATVSSPLTSPTS 279
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
943 LEGSSALN-----VLNANGKELGRTITITLKSPDWIKTL--EEMSLKINVLPSSTS 995
QY 280 PSTLSKSESVSATEDLAPDAQEDNSIKELLREERIEKEGSEASSSE---EDDPL 336
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
996 STILL---GEDAEVTADFMEGDV--DDYSAEVEIL-PQLQPSSSSALARPVLHPGPV 1049
QY 337 PACNGPAQAQPSPTTERAKSOEVLPS---STTPSGALSPSQSPSSATEVVLRTTA 393
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1050 PASHLFYRRGVPVPSL-----PVRPSRAPSRTPGP-----PASQ---SSPVDTLPATQLQ 1095
QY 394 SEGSEQ-----PKRASIQRTSAPSPRPPPRATASPRPSSGNIPSSPTASG-----440
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1096 QKDSQTLKPPPPPPPPVAPAPAPAPQPPPPSGARSPAPAKERVWSTRKAQERPRD 1155
QY 441 --GGSPTSRSL-CTGTASPRTSLEVPNPEPPEKFPVTPPEAKENENIHQNP-----492
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1156 NLGSQLPPQGLPGPLAGHSAARPIIP---PRAGVISAPESHGRVSAAGRLTPESQRT 1212
QY 493 -ELCTSPILMTS-----QVASEPGEAKMEDKEKDKNLISADSEGGDQLOQVSMVPENNLL 547
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1213 XEVLKGPALLPEPLKQKQALVPVPSLAPPSPQEMOEPLIAVAFLAQSALQPSLE-----1266
QY 548 TAPEPQEEVSTSEN 561
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1267 TPPQPPPRSRSSHS 1280

```

Search completed: March 4, 2004, 17:28:20  
Job time : 17.3694 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:29:55 ; Search time 43.1083 Seconds  
(without alignments)  
2762.598 Million cell updates/sec

Title: US-10-069-540A-2

Perfect score: 2855

Sequence: 1 MAEGKAGGAGLFAKQVQK.....NNLTAPPEQBEVSTSENPL 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	829	29.0	482	15	US-10-116-275-238
2	816.5	28.6	434	9	Sequence 238, App
3	793.5	27.8	451	14	Sequence 22, Appl
4	675	23.6	404	9	Sequence 4, Appl
5	238	8.3	489	15	US-09-879-957-24
6	230	8.1	452	15	US-10-369-493-2512
7	206.5	7.2	462	10	US-10-369-493-3119
8	203	7.1	1367	9	US-09-819-104A-5
9	202.5	7.1	3507	15	US-09-801-368-108
10	200.5	7.0	3913	15	US-09-369-493-5784
11	199.5	7.0	1297	15	US-10-334-143-45
12	199	7.0	284	15	US-09-291-417-14
13	199	7.0	818	15	US-10-369-493-2410
14	197.5	6.9	6542	15	US-10-104-047-2546
15	197	6.9	1360	9	US-10-369-493-5013
					Sequence 2, Appl

16	197	6.9	1360	14	US-10-355-975-14	Sequence 14, Appl
17	195	6.8	600	15	US-10-357-587-16	Sequence 16, Appl
18	194.5	6.8	2441	13	US-10-109-886-8	Sequence 8, Appl
19	192	6.7	758	15	US-10-094-749-2451	Sequence 4, Appl
20	191.5	6.7	1848	10	US-09-839-996-6	Sequence 6, Appl
21	191.5	6.7	1848	14	US-10-080-505-6	Sequence 6, Appl
22	191	6.7	742	14	US-10-308-448-11	Sequence 11, Appl
23	190.5	6.7	2665	9	US-09-864-761-34248	Sequence 34248, A
24	190.5	6.7	3564	14	US-10-177-293-423	Sequence 423, App
25	190	6.7	742	14	US-10-203-860-2	Sequence 2, Appl
26	190	6.7	742	15	US-10-341-434-85	Sequence 85, Appl
27	189	6.6	1268	15	US-10-353-690-122	Sequence 122, App
28	188	6.6	474	14	US-10-032-585-7847	Sequence 7847, Ap
29	188	6.6	742	14	US-10-203-860-4	Sequence 4, Appl
30	187.5	6.6	528	10	US-09-840-746-20	Sequence 20, Appl
31	187.5	6.6	744	10	US-09-769-787-184	Sequence 184, App
32	187	6.5	258	15	US-10-104-047-3034	Sequence 3034, App
33	186.5	6.5	2440	15	US-10-341-434-236	Sequence 236, App
34	186	6.5	2478	9	US-09-815-242-5816	Sequence 5816, Ap
35	186	6.5	2478	9	US-09-815-242-12967	Sequence 12967, A
36	184.5	6.5	455	14	US-10-032-585-7399	Sequence 7399, Ap
37	184.5	6.5	511	9	US-09-864-761-34590	Sequence 34590, A
38	184.5	6.5	545	14	US-10-029-386-32280	Sequence 32280, A
39	184.5	6.5	551	9	US-09-864-761-34727	Sequence 34727, A
40	184.5	6.5	1702	10	US-09-839-996-5	Sequence 5, Appl
41	184.5	6.5	1702	14	US-10-100-957A-152	Sequence 152, App
42	184	6.4	1125	14	US-10-100-957A-22	Sequence 22, Appl
43	184	6.4	1610	14	US-10-032-585-7122	Sequence 7122, Ap
44	182	6.4	841	14	US-10-369-493-5455	Sequence 5455, Ap
45	181.5	6.4	1974	15	US-10-369-493-5455	Sequence 5455, Ap

#### ALIGNMENTS

#### RESULT 1

US-10-116-275-238  
; Sequence 238, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 238  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-238

Query Match 29.0%; Score 829; DB 15; Length 482;  
Best Local Similarity 39.2%; Pred. No. 1.2e-34;  
Matches 211; Conservative 70; Mismatches 167; Indels 90; Gaps 16;  
QY 1 MAE-GKAGGAGLFAKQVQKFSQAQKVLQKLGKAVETKDERPEOSASNFYQQQAEHKK 59  
Db 1 MAEMSGKVTAGKTASNVQKLTQAQKVLQKLGKADETKDEQEQCVQNFNQLTEGTR 60  
QY 60 LYKDLKFLSAVKYWHSSKEVSTLQEIYSSWDGHEELKATVWNNLLWEDYEKLD 119  
Db 61 LQDURTTTSLASVKANHEASKKLNELQEVPEPDMFGDEANKTAENNDLLWMDYHOKLVD 120  
QY 120 QAVRTMEIYVQAFSEIKERIAKRGKLVVDYDSARHLEAVQNA-KKDEAKTAKAEFEFNK 178

Db 121 QALLTMDTYLGQPPDIKSRKGRKLVYDSARHHYSLQAKKQKAEELIK 180  
Qy 179 AQTVFEDLNOELLELPILYNSRIGCVYTFQNIENLNDVYREMSKLNHLYVMKLE 238  
Db 181 AQVFEEMVNDLQELPFLSNRSGVGYNTFQSIAGLEENFHKMSKLNQNLNDVLVLE 240  
Qy 239 KQHSNKVFFVVKGLSSRRSLVSPVRVATVSSPLTSPSTLSLSESESVSATEDL 298  
Db 241 KQSGNTFTVKAQPSDN-----APAKNGKSPSPDGSPATPEIRVNHPEPAGGA--- 291  
Qy 299 APDAAGDENSETKELLEEEIEKEGSEASSEDDPLPACNGPAQAQSPSTTERAKSQE 358  
Db 292 TPGATLPKSPQLRK-----GPPVPPPKHTPSKEVKQE 325  
Qy 359 EVLP-----SSTTSPGALSPSPQSSSATEVLRTRTASGSEOPKKRASIQR 408  
Db 326 QILSLFEDFVPEISVTTFS-----QP-AEASEVAGGTQPA-A-GAEPFGTAA--- 371  
Qy 409 TSAPPSPRPPTATASPRPSSGNIPSPPTASGGSPPTPRASLGTG-----TAS 457  
Db 372 SEASSSLPAVVVETFPATVNGTVEG---GSGAG-----RLDLPFGFMKVQAQHDYAT 423  
Qy 458 PRTSLVSNPPEPEKVPETPEAKE-----NENIHQNP-E-LCTS--PLTMSQV 505  
Db 424 DTDELQKAGDVVLVTFQNPEDQEGWLMGVKESDMNQHLEKRCRGVFPENFETRV 481

## RESULT 2

US-09-879-957-22  
; Sequence 22, Application US/09879957  
; Patent No. US20020034755A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; HOFFMAN, No. US20020034755A1h  
; KAY, Brian K.  
; FOMKES, Dana M.  
; MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; USING SAME

NUMBER OF SEQUENCES: 227  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/879,957  
FILING DATE: 13-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,915  
FILING DATE: 03-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>

; TOPOLOGY: unknown  
; MOLECULE TYPE: Peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-879-957-22  
Query Match 28.6%; Score 816.5; DB 9; Length 434;  
Best Local Similarity 45.1%; Pred. No. 4.4e-34;  
Matches 193; Conservative 47; Mismatches 105; Indels 83; Gaps 11;  
Qy 1 MAE-CKAGAGLFAKOVKKFSAQKVLQKLGKAVETKDERPEQASNFYQQQAEHGK 59  
Db 1 MAEMSGKGVTAGKIASNVQKLTTRAQKVKLGKADETKDEQEQCVQNFNKLTEGTR 60  
Qy 60 LYKDLNLFSAVKYMHSSKESVSETLQEIYSSSEWDGHEELKAIWMNDLLWEDYEKGLAD 119  
Db 61 LQKDLRTVLASVKAMHEASKGLSECLQEVYEPWPGRDEANKIAENNDLLWMDYHQLVD 120  
Qy 120 QAVRWELVYVQAFSEIKERIAKRGKLVYDSARHHLEAVQNA-KDEAKTAKAESEFNK 178  
Db 121 QALLTMDTYLGQPPDIKSRKGRKLVYDSARHHYSLQAKKQKAEELIK 180  
Qy 179 AQTVFEDLNOELLELPILYNSRIGCVYTFQNIENLNDVYREMSKLNHLYVMKLE 238  
Db 181 AQVFEEMVNDLQELPFLSNRSGVGYNTFQSIAGLEENFHKMSKLNQNLNDVLVLE 240  
Qy 239 KQHSNKVFFVVKGLSSRRSLVSPVRVATVSSPLTSPSTLSLSESESVSATEDL 298  
Db 241 KQSGNTFTVKAQPSD-----NAPEKNGKSPSP----- 268  
Qy 299 APDAAGDENSETKELLEEEIEKEGSEASSE---EDDPLPACNGPAQAQSPSTTERAK 355  
Db 292 TPGATLPKSPQLRK-----GSPATPEIRVNHPEPA-----SGASPGATIPKSP 302  
Qy 356 SQEVLPSSTTSPGALSPSPQSSSATE-----VVLRTRTAS-----EGSQPKKRAS 405  
Db 303 SQ-----PAEASEVGAQEPGTAASEATSSSLPAVVVETFSATVNGAVEGS-----AG 352  
Qy 406 IQTSAPP 413  
Db 353 TGRLLDLP 360

## RESULT 3

US-10-123-807-4  
; Sequence 4, Application US/10123807  
; Publication No. US20030166021A1  
; GENERAL INFORMATION:  
; APPLICANT: Wistar Institute of Anatomy & Biology  
; Prendergast, George C.  
; Sakamuro, Daitoku  
; TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein  
; Compositions and Uses Therefor  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/123,807  
; FILING DATE: 16-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/445,247  
; FILING DATE: 03-Dec-1999  
; APPLICATION NUMBER: US 08/870,126



FILING DATE: 06-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: WS\*60DPCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 451 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-123-807-4

Query Match 27.8%; Score 793.5; DB 14; Length 451;

Best Local Similarity 36.6%; Pred. No. 6.9e-33;

Matches 193; Conservative 66; Mismatches 152; Indels 117; Gaps 11;

QY 10 AGLFAQVQKPSRAQEKVLQKLGKAVETKDRRFRQASNFYQQAGHKLKYLKDNFLS 69

DB 8 AKGIASNVQKLTQAEQKVLQKLGKADETQDEQPEQCVQNFKNQLTETRLQKDLRTYLA 67

QY 70 AVKVMHESKRVSETLQETYSSEWDGHEELKAIWVNDLLWEDYEKLADQAVRTMEIYV 129

DB 68 SVKAMHEASKKNECLQEVFEDWDFGRDEANKIAENNDLLWMDYHQKLVQALLTMDTYL 127

QY 130 AQPSEIKERIAKGRKLVYDYSARHLEAVQNA-KDEAKTAKAEFEFNKAQTVFEDLNQ 188

DB 128 GQPPDIKRIAKGRKLVYDYSARHHYESLQPAKKDEAKTAKAEELIKAKQVFEEMNV 187

QY 189 ELLEELPILYNSRIGCVYTFQNIENLRDVFYREMSKLNHLYEVMKLEKQHSNKVFV 248

DB 188 DLQEEPLSNRKGVGYVNTFQSIAGLEENFKEMSKLNQNLNDVLVGLERQHGSENTFTV 247

QY 249 KGLSSSRSSRLVSPVTRATVSSPLTSPSTSLKSESESVSATDLAPDAQGEDN 308

DB 248 K-----AQPRKSKLSRLRRKNSDNAPAKGNKSP----- 278

QY 309 SEIKELLEEEIEKGESEASSEEDDPLPACNGPAQAQPSPTTERAKSQEVLSSSTPS 368

DB 279 -----SPDGSFAATPEIRVNHE-----PE 298

QY 369 PGALSPSGQPSSSATEVVLRTTASEGSEQPKKASQRTSAPSPRPPRATASPRPS 428

DB 299 PAGATPGA-----TLPKSPQAEASEVAGGTQPAAGAQEFCEETAASEAA 344

QY 429 SGNIPS-----SPTASGGSPSPRASLGTTG-----TASPRTSLEVSFN 467

DB 345 SSSLPAVVVETFPATVNGTVGGSG--AGRLDLPFGFMFKVQAQHDYATDTDELQALQAK 402

QY 468 PEPEKPVRTPEAKE-----NENIHQNPE-ELCTS--PTLMTSQV 505

DB 403 DVVLVTFQNPPEQDEGLWLVGKESDMNQHKLEKRCRGVFPENFTERV 450

# RESULT 4

US-09-879-957-24

Sequence 24, Application US/09879957

Patent No. US20020034755A1

GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.

HOFFMAN, No. US20020034755A1h

KAY, Brian K.

FOULKES, Dana M.

MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/879,957  
FILING DATE: 13-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,915  
FILING DATE: 03-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-174  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 404 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 24:

US-09-879-957-24

Query Match 23.6%; Score 675; DB 9; Length 404;

Best Local Similarity 34.1%; Pred. No. 6.5e-27;

Matches 169; Conservative 62; Mismatches 133; Indels 132; Gaps 12;

QY 42 RFEQASAFYQQAAGHKLKYLKDNFLSAVKVMHESKRVSETLQETYSSEWDGHEELKA 101

DB 8 RFEQCVQNFQKLTETGLRQDLRTYLA SVKAMHEASKKNECLQEVFEDWDFGRDEANK 67

QY 102 IYVNDLLWEDYEKLADQAVRTMEIYVVAQFSEIKERIAKGRKLVYDYSARHLEAVQN 161

DB 68 IAENNDLLWMDYHQKLVQALLTMDTYLQGPDIKRIAKGRKLVYDYSARHHYESLOT 127

QY 162 A-KDEAKTAKAEFEFNKAQTVFEDLNQELLELPILYNSRIGCVYTFQNIENLRDVFY 220

DB 128 AKKDEAKTAKAEELIKAKQVPEEMNVDLQEEPLPSLWNRKVGIVNTFQSIAGLEENFH 187

QY 221 REMSKLNHLYEVMKLEKQHSNKVFVVKGLSSSRSSRLVSPVTRATVSSPLTSPSTSP 280

DB 188 KEMSKLNQNLNDVLVGLERQHGSENTSTVK----- 216

QY 281 STLKSESESVSATDLAPDAQGEDNSEIKELLEEEIEKGESEASSEEDDPLPACN 340

DB 217 -----AQPSDNAPAK-----GNKSPS----- 232

QY 341 GPAQAQPSPTTERAKSQEVLPSSTTPSPGGALSPSGQPSSSATEVVLRTTRTASEGSEQ 400

DB 233 -PPDGSFAATPEIRVNHE-----PEPAGATPGA-----TLPKSPSQ 269

QY 401 KKRASQRTSAPSPRPPRATASPRSSGNIPS-----SPTASGGSPSPSPRAS 450

DB 270 ABASEVAGGTQPAAGAQEPGETAASEASSSLPAVVVETFPATVNGTVGGSG--AGRLD 327

QY 451 LGTG-----TASPRTSLEVSFNPEPEKPVRTPEAKE-----NENIHQNPE 492

DB 328 LPPGFMFKVQAQHDYATDTDELQALQAGDVVLVTFQNPPEQDEGLWLVGKESDMNQHK 387

QY 493 -ELCTS--PTLMTSQV 505

RESULT 6  
 US-10-369-493-3119  
 ; Sequence 3119, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 3119  
 ; LENGTH: 452  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 ; US-10-369-493-3119  
 Query Match 8.1%; Score 230; DB 15; Length 452;  
 Best Local Similarity 21.5%; Pred. No. 0.00033;  
 Matches 96; Conservative 68; Mismatches 188; Indels 94; Gaps 14;  
 QY 23 RAQKVLQKLGKAVETKDERFEQASNFYQQQAEGHKLKDLKFLSAVKVMHESKRV 82  
 DB 1 QAPQKQKFNGLGHTKOPVYIDSRERFQETETKRLHDESKKYFEALINGMLHQIEFS 60  
 QY 83 ETLOEYIS-----SEWDG-----HEELKALVWNNLLWEDYEKLADQ----- 120  
 DB 61 KAMTEIYKPISGRMSDDPSLVPHGNVEGIAACEYEAVV-----KDLQETLAPLEMI 113  
 QY 121 ---AVRTMEIYVAQSEIKERIAKGRKLVYDSARHLEAVQNAK-----KDEAKTAAE 173  
 DB 114 EARVIRPANELLDVIVIRKTAVERHKKLDIDRHRATLKLQDKKRSADKAMWKA 173  
 QY 174 EEFNKATQVFDLNOELLELPILYNSRIGCVYTFQINSLNR-DVFYREMSKLNH---N 229  
 DB 174 NEVQATQYVYNDLLKDELKPLALERQFQIPLFQSFYFQMLNIFYTLHEKMQHCDIG 233  
 QY 230 LYEVMSKLEKQHNKVFVVKGLSSSRSLVSPVVRTATVSSPLTSPSTLSKSES 289  
 DB 234 YFDLTLDIE-----EAFYAKRGDQVRAEALSIVKFTTKMRP----- 272  
 QY 290 ESVGATEDLAPDAAGDENSEIKELLEEBEIEKEGSEASSEEDDPLPACNPAQAPSP 349  
 DB 273 -----PKYQRPGLAEGN-----KPAGLLTAGSSTTTS-----TGPEAAPS 309  
 QY 350 TTERAKSQEVLPSSTTSP-----GGALSPGQSPSSATEVVLRT--RTASEGSEOPKKR 403  
 DB 310 TAPREWEQQQAPAVTAPRFRWETOAEVHAHPPPVSVKPSLTHATPLATPLQASAPK 369  
 QY 404 ASIQTSAPPSPRPPPRATSPRSS 429  
 DB 370 SLSMAAAKAKPPPK-----PKPKA 391

RESULT 7  
 US-09-819-104A-5  
 ; Sequence 5, Application US/09819104A  
 ; Publication No. US20030027137A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chen, J. Don  
 ; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
 ; TITLE OF INVENTION: AND USES THEREFOR  
 ; FILE REFERENCE: UMG-030  
 ; CURRENT APPLICATION NUMBER: US/09/819,104A  
 ; CURRENT FILING DATE: 2001-03-27  
 ; PRIOR APPLICATION NUMBER: 60/193,138

DB 388 LEKRGVFPENFTERV 403  
 RESULT 5  
 US-10-369-493-2512  
 ; Sequence 2512, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 2512  
 ; LENGTH: 489  
 ; TYPE: PRT  
 ; ORGANISM: Schizosaccharomyces pombe  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(489)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-369-493-2512  
 Query Match 8.3%; Score 238; DB 15; Length 489;  
 Best Local Similarity 22.08%; Pred. No. 0.00014;  
 Matches 110; Conservative 69; Mismatches 194; Indels 128; Gaps 16;  
 QY 15 KOVKKFSRAQKVLQKLGKAVETKDERFEQASNFYQQQAEGHKLKDLKFLSAVKVM 74  
 DB 4 KGFTKALARTQTLSRKNFVNGEITKDIYEDAGRPFKSLTEAKGAEDAKKTTDANG 63  
 QY 75 HESKRVSETLOEYI-----SSEWDG-----EELKALVWNNLLWEDYEK 116  
 DB 64 LNHQIGFADACIEIYKPISGRASDPESVEQEGNAEGIAEAYKEIVY-----DLQKN 116  
 QY 117 LADQAVRTMEIYVAQF-----SEIKERIAKGRKLVYDSARHLEAVQNAK-- 163  
 DB 117 LASE-----MDVINTRIVNPTGELLKIVKVDKLLKRDHKOLDYDRHRSSFKLQSKDK 172  
 QY 164 --KDEAKTAAEBEENKACQTFEDLNOELLELPILYNSRIGCVYTFQINSLNRDVFYR 221  
 DB 173 SLKDEKGLYEATAFEQSSQSEYTYNEMLKEELPKLALAQSFIAPIFGQ-----FY 225  
 QY 222 ENSKLNHLNLYEVMSKLEKQHNKVFVVKGLSSSRSLVSPVVRTATVSSPLTSPSTPS 281  
 DB 226 MQLNVYVLYKMSHCEIQFD--FNTDILESYYERRRGDVKDRAEALITIKFTAKPTYK 283  
 QY 282 TSLKSESESVSATEDLAPDAAGDENSEIKELLEEBEIEKEGSEASSEEDDPLPACNG 341  
 DB 284 RPKMGPGGKDATAS-----SSSFSSKREE----- 308  
 QY 342 PAQAPSPPTTERAKSQEVLPSSTTSPGGLSPGQSPSSATEVVLRTTRTASEGSPK 401  
 DB 309 -AAAESSST-----ATDIPPPYSTPSVAGSDYS--TPSAGYQVVTQTTTTTAAQAQ 360  
 QY 402 KRASIORTSAPPSPRPPPRATSPRSSGNISSPTASGGSGFTSPRASLGTGTASPRTS 461  
 DB 361 PQAAF-----PPPP-----VMPQFAA-----AAVTPVZAPVAA 389  
 QY 462 LEVS--PNEPPEKPVRTPEAK 481  
 DB 390 AAAAVPVPPPPAPAPAAPAAE 410

; PRIOR FILING DATE: 2000-03-29  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 2462  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-819-104A-5

Query Match 7.2%; Score 206.5; DB 10; Length 2462;  
 Best Local Similarity 19.6%; Pred. No. 0.039;  
 Matches 113; Conservative 91; Mismatches 213; Indels 159; Gaps 22;  
 QY 15 KOVKKPRAQKVLQKLGKAVETKDERPEQASNF-----YQQQAEQ 57  
 DB 426 RQVWNWSE-QER-----DTREKFMQHPKNGFLIASFLERKTVASCVIYYITKK 475  
 QY 58 HKLYKDL--KNFLSAVKVMHSSKRVSETLQEI--YSEWDGHEELKAIVWNNOLLMEDYE 114  
 DB 476 NENYKSLVRSYRRGKSKQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 535  
 QY 115 EKLDAQVTRMIYVAQSEIKERIAKGRKLVYDSAR-----HLEAVONA 162  
 DB 536 NEKZELSKETDITSGEDNDEKAEVASKRKTANSQGRKGRITRSMANANHEETATPQ 595  
 QY 163 KKDEAKTAKABEEFNKAQTVFEDLNQELLE-----LPILYNGRIGCYVTFONISNR 216  
 DB 596 QSEELASMEMNNESSRWTEEMETAKGLLEHGRNWSAIAARMVSK-----TVSOCK 646  
 QY 217 DVF--YREMSKLNHLNLYEVMKLEKQHNKVFVVKGLSSRRRLVIVPPVTRATVSSPL 274  
 DB 647 NFVFNKQKQNLDELLOQKLMKEKER-----NARRK----- 678  
 QY 275 TSPTSTLSKSESVSATEDLAPDAQGEDNSEIKELLEEEIEKEGSEASSSEDD 334  
 DB 679 -----KKKTFAASEETATPPAAEDDEMGASANEELAEAEASQASGNEV 727  
 QY 335 P-LPACNGPAQPSPTTTERAKSQEVLPSPTTTPGGALSPGQSPSSSATEVVLRTTA 393  
 DB 728 PRVGCSGPAANVNSDTE-----SVPSFRSEATKDTGPKTGTGTE----- 767  
 QY 394 SEGSEQPKKASRTQTSAPPSRPP--PPRATASPPSSGNIPSSPTASGGSGSPTRASL 451  
 DB 768 -----ALPAATQPPVPPPPPEAPAAAPSP-----VPDASG-----PSPEPS- 805  
 QY 452 GTGTASPTLSVSPNPEPEKPVRTPEAKEN-----ENIHNONPEELCTSTPLMTSQVAS 507  
 DB 806 -PSPAPPATVDKQEAAPAPAPQTEDAKQKSEAEEIDVGKPEE-----PEASE 855  
 QY 508 EPGEAKQMEDKENDKLISADSSEGGQDQLO-Q-VMVP 542  
 DB 856 EPPESVKSDHKEETEE-EPEDKAKGTEAETVSEAP 890

RESULT 8

US-09-801-368-108  
 ; Sequence 108, Application US/09801368  
 ; Patent No. US20020128250A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Busby, Robert  
 ; APPLICANT: Cali, Brian  
 ; APPLICANT: Hecht, Peter  
 ; APPLICANT: Holtzman, Doug  
 ; APPLICANT: Madden, Kevin  
 ; APPLICANT: Maxon, Mary  
 ; APPLICANT: Milne, Todd  
 ; APPLICANT: No. US20020128250A1man, Thea  
 ; APPLICANT: Royer, John  
 ; APPLICANT: Salama, Sofie  
 ; APPLICANT: Sherman, Amir  
 ; APPLICANT: Silva, Jeff  
 ; APPLICANT: Summers, Eric  
 ; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

; FILE REFERENCE: 109272.147  
 ; CURRENT APPLICATION NUMBER: US/09/801,368  
 ; CURRENT FILING DATE: 2001-03-07  
 ; PRIOR APPLICATION NUMBER: US 09/487,558  
 ; PRIOR FILING DATE: 2000-01-19  
 ; PRIOR APPLICATION NUMBER: US 60/160,587  
 ; PRIOR FILING DATE: 1999-10-20  
 ; NUMBER OF SEQ ID NOS: 440  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 108  
 ; LENGTH: 1367  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-801-368-108

Query Match 7.1%; Score 203; DB 9; Length 1367;  
 Best Local Similarity 25.2%; Pred. No. 0.03;  
 Matches 82; Conservative 51; Mismatches 155; Indels 38; Gaps 11;  
 QY 252 SSSRRRLVIVPPVTRATV---SSPLTSPSPSTLSKSESVSATEDLAPDAQGEDN 308  
 DB 360 SSSTTES--SSAPVTSSSTSSAPVTSSSTSS--SAPVPTSSSTSSAPVTSSSTTE 416  
 QY 309 SEIKELLEEEIEKEGSEASSSEDDPLPACNGPAQAO----PSPITERAKSQEVLPS 364  
 DB 417 SSSAPVTSSSTSSAPVTSSSTSSAPVTSSSTSSAPVTSSSTSSAPVTSS 476  
 QY 365 TTPSPGGALSPGQSPSSSATE---VVLRTASGEGSEKPKKASRTQTSAPPSRPP 420  
 DB 477 TTES--SSAPVTSSSTSSAPVTSSSTSSAPVTSSSTSSAPVTSSSAPAPTSS 533  
 QY 421 ATASPPRSGNIPSPASGGSGTSPRASLGCTGCTASPTSLSVSPNPEPEKPVRTPE 480  
 DB 534 TTES--SSAPVTSSSTSSAPVTSSSTSSSTTPVT---SSTSSSAPVTSS 586  
 QY 481 KENENIHNONPEELCTSTPLMTSQVASPEGSEAKMEDKENDKLISADSSEGGQDQVSM 540  
 DB 587 STTESAPVP-----TPSSSTSSSAPAPTSSSTSSAPVTSSSTSS 637  
 QY 541 VPENNNLT-----APEPOEVSTSEN 561  
 DB 638 VPTSSSTSSSAPVTSSSTSS 663

RESULT 9

US-10-369-493-5784  
 ; Sequence 5784, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 5784  
 ; LENGTH: 3507  
 ; TYPE: PRT  
 ; ORGANISM: Caenorhabditis elegans  
 US-10-369-493-5784

Query Match 7.1%; Score 202.5; DB 15; Length 3507;  
 Best Local Similarity 25.5%; Pred. No. 0.096;  
 Matches 99; Conservative 51; Mismatches 175; Indels 63; Gaps 15;  
 QY 233 VMSKLEKQHNKVFVVKGLSSSS---RRSLVISPVPVTRATVSSPLTS---PTSPST-LSL 285

Db 2066 VKSTPKESSSEITVK-ISSKQPEVTESSVSSPSTPTTQSVTSIVPETSSTVLS 2124  
Qy 286 KSESSEYATE-----DLAPDAQOEDNSEIKELLEBEIEKEGSEASSSEEDPL 336  
Db 2125 EAPVTSSTPTEVHTSSSTKPSLSASSTTGTNSTPSTSSLASVSTSAPEGTSASVAP 2184  
Qy 337 PACN-GPAQAQPSPTTERRAKSQEVLPSSTTPSPGGALSPSGOPSSSATEVVL----- 388  
Db 2185 KLSLSFDVQPSKTDFATESVQASET--SSGTSVASTSEPSHVTKLSITSNPS 2242  
Qy 389 -----RRTASEGSEQPKKASQRTASPPRPPPPRATASPPRPPSPPTASG 440  
Db 2243 SVPTSPKSTPTPESTEQPTTTPSQSLTPMNSSEVLTTSEPHVLSSSL--SPDVS- 2299  
Qy 441 GGSPTSFRASLGCTASPTSLSVSPNPEP---EKPV-----RTPEA 480  
Db 2300 QSSTTPNLSSTVETPKTSSEVSLNSEPSTTEAPTTLSPDILSTTNNLSQSTVST 2359  
Qy 481 KENENIHQNPPEELCTPTLMTSQVASEPGEAKM--EDKXKNKLISADSS--EQDQL 536  
Db 2360 EDRSEISSENSEKPTSAPELVTSVTHVASSPDVPTESSEPDLTGSSSTENIPEASSKQ 2419  
Qy 537 QVSMVPENNLTAPPEQSEVSTSENPOL 564  
Db 2420 TISSTPTPTTASE-EPTKSTGMSDPL 2446

RESULT 10  
US-10-334-143-45  
; Sequence 45, Application US/10334143  
; Publication No. US20040009549A1  
; GENERAL INFORMATION:  
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1543  
; CURRENT APPLICATION NUMBER: US/10/334,143  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: 60/343,169  
; PRIOR FILING DATE: 2001-12-31  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 45  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-334-143-45

Query Match 7.0%; Score 200.5; DB 15; Length 3913;  
Best Local Similarity 20.7%; Pred. No. 0.14; Indels 211; Gaps 33;  
Matches 144; Conservative 99; Mismatches 240; Indels 211; Gaps 33;  
Qy 12 LFAQVQKK-FSRAQKV-LQKLGKAVETKDE-----RFEQSA-- 47  
Db 2297 VFAREKQKQKADLPDESVCQDFMWLTKTDEHAQSNVVDSDGSDNVKQRTENSSKA 2356  
Qy 48 --SNFYQQAQGHKLYKDKNFLSA-----VKYWH--ESS-----KRYSETL 85  
Db 2357 MPDSFSEQA-----KDLACHITSDDLATRGFMWKKVFRTWSSGATNKSQKEKLSHVL 2410  
Qy 86 QEIVSSEWDGHEELKAIWVNDLL--WEDYBEKLDAQVRTWETVYVAQFSIKERIAKGR 144  
Db 2411 VHDVRENHIGPESKSVQDKNEFMSVTERERKLTN-----GSLSEIKEMTVKSPS 2461  
Qy 145 KLVDYSARHLEAVQNAKDEAKTAKAEFEFNKAQTFEDLQELLEEPLILYNSRIGC 204  
Db 2462 KKVLV---REVY-----VKEGDHPGGLDQPSRES-----SAVSHIPV---RVAD 2502  
Qy 205 YVTIFQNLISNLRDVFYRMSKLNENLYEVMK-----LEKQSNKVFVVKGLSS----- 253  
Db 2503 ERMLS--SNIPDGFCEQASAPPKHLSQKLSQSSSKETVETQHFNSIEDEKVTYSEISK 2560

Qy 254 -SSRSLV-ISPVRTATVSSPLTSPT-SPSTLSLKSESES-----VSATEDLAP 300  
Db 2561 VSKHSGYVGLCPLEBETE-----TSPTKSPDLSFSPGKESPSDVFHDSPIDGLEKLAP 2615  
Qy 301 DAAQGEDNSSEIKEL-----LEBEEI-----E 321  
Db 2616 -LAQTEGGKEIKTLFVYVSVQVQKYEKEIQQGVKKIISQECKTVQETRGTYTTRQQ 2674  
Qy 322 KEGSEASSSEEDDPLPACNGPAQAQPSPTTERRAKSQEVLPSSTTPSPGGALS-PSQPS 380  
Db 2675 KQPPSPQSPEDDTLEQVSLDSSGKSPLTPTETPSSEVSYFTSKTPDSLIAIIPCKPS 2734  
Qy 381 SSATEVVLTRTASGSGOPK-----KRASIORTSAPPSPRPPPPRATASPPSSGNI----- 432  
Db 2735 P-----IPEVSESESEEEQAKSTSLKQTTVEBTAVEREMPNDVSKDNQRPKNRVAYIE 2789  
Qy 433 -PSSPTASGGSPTRASLGTTGATSPRTSLEV-----SPNPE 469  
Db 2790 FPPPPPLDADQIESDKHY-----LPEKVDMEVNLQDEHDKVQLAEPVIRVQPPSPV 2844  
Qy 470 PPEKPVRTPEAKENINHNQNPPEELCTPTLMTSQVASEPGEAKMEDKXKNKLISADS 529  
Db 2845 PPGADV--SDSSDSESIYQPVVKKYTF-----KLKEVDDEQEKPKASAEK 2889  
Qy 530 SEGQDOLQVSMVPENN--NLTAPEPEQSEVSTSEN 561  
Db 2890 ASNQKESLNGSGKONEPGLGLDPSQNELAQNGN 2923

RESULT 11  
US-09-291-417-14  
; Sequence 14, Application US/09291417A  
; Publication No. US20030050230A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHYTE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 240/300  
; CURRENT APPLICATION NUMBER: US/09/291,417A  
; CURRENT FILING DATE: 1999-04-13  
; EARLIER APPLICATION NUMBER: US 60/081,784  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 147  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 1297  
; TYPE: PRT  
; ORGANISM: Mammalian (Human) 2C2  
US-09-291-417-14

Query Match 7.0%; Score 199.5; DB 10; Length 1297;  
Best Local Similarity 18.6%; Pred. No. 0.042;  
Matches 113; Conservative 113; Mismatches 258; Indels 123; Gaps 21;  
Qy 15 KQVQKFSRAQKVQKLGKAVETKDEPERFEQASAFYQQAQGHK-----LYK 62  
Db 262 ROVRIQLKDHIDRTKKRGEKDETE---YEYSGSEEEENDSGEPSSILNLPRESTLRR 318  
Qy 63 DLKNFLSAVKVMHESKRVSETLQEIYSSSEWDGHEELKAIWVNDLLWEDYEE---KLAD 119  
Db 319 DFLQLQANKERSEALRR-----QQLSQQORENEHEKQLLAERQKRIEEOKEQRRLEE 373  
Qy 120 QAVTMEIYVAQFSEIKERIAKGRKLVVDYSARHLEAVQNAKDEAKTAKAEFEFNKA 179  
Db 374 QORREKLRQQRQREQRHYEQMREERREARAEHQYKRRQLBEQQAERLQRLQKE 433  
Qy 180 QTVFEDLNQELLELPI-----LYNSRIGCVITFQNLISNLRDVFYRMSKLN----- 227  
Db 434 RDLVLSLQHQQRQRPVKKKPLHYHKEG--MSPEKPAWAKEV--EERSLNRQSSPAMP 489  
Qy 228 HNLVEVMKLEKQHSNKFVVKGLSSSSRRSLVISPPVRTATVSSPLTSPTSPSTLSLKS 287

[illegible]

```

RESULT 12
US-10-369-493-2410
; Sequence 2410, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2410
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(284)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2410

```

Query Match	7.0%;	Score 199;	DB 15;	Length 284;
Best Local Similarity	26.4%;	Pred. No. 0.0073;		
Matches	60;	Conservative 42;	Mismatches 105;	Indels 20; Gaps 6;
Qy	28	V L Q L G K A V E T K D E F E Q S A N F Y Q Q A E G H K L Y K D L K N F L S A V K V M H E S S K R V S E T L Q E	87	
Db	37	V M K T G H V E R T V D R E F E T E E R Y R T M S A K K Q G A K Y D L A R A N T A S Q T R I A N T I D A	96	
Qy	88	I Y S S E W D G H E E L K A I V W N D L L W E D Y E E - - K L A D Q A V R T M E I - - - - Y V A Q F S I K E R I	139	
Db	97	F Y G D A - G S K O G V S A Y R Q - - V V E D L D A D T V K E L D G P F R T V L D P I R S F C S Y F P D I N A A I	152	
Qy	140	A K E G K L V D Y D S A R H L E - A V Q N A K D E A K T A S E E F N K A Q T V P E D I N Q E L L E B L P I L Y	198	
Db	153	T R N H K L L D H D A M R A K V Q K L V D K S N P T T K L P T E K A A M A K E V E T L N Q N L V S E L P O L I	212	
Qy	199	N S R I G C Y V T I F O N I S N L R D V F R E M S K N H N L Y E V M S K L E K O H S K V	245	

```
DB      213 ALRPVYLDPSFEALVKIQLRFREG-----YEQAAQVQQYFDNSV 252

RESULT 13
US-10-104-047-2546
; Sequence 2546 Application US/10104047
; Publication NO. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392a1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2546
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2546
```

```

Query Match      7.0%; Score 199; DB 15; Length 818;
Best Local Similarity 21.5%; Pridat.No.0.026;
Matches 128; Conservative 75; Mismatches 231; Indels 158; Gaps 24;

QY   23 RAQEKVLQKLGAVETKDEFEQSANFYCQAEGHKLY-----XDLKNFLSAYKVNM 74
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   120 KMQERMSAQLA-AAESRQKKLEMKLQALQEQEHKLAARLEBERGNGKWVLMKYEC 178
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY   75 HESSKRVSSETLOBIYSSEWDGHEELKAIWVNNDLLMEDYEKKLADQAVTWEI---VVAQ 131
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   179 KQLSKGVIEBAQKL-----EDVMALKEEKKTKTWLEEELSAAE 216
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   132 FSEIKERTIAKGRKLVLDYDSARHHLAEVONAKDKAKTAKEEFENFKAQTVPFDNLQELL 191
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   217 XRRSTEMAQOMKQLSEFDERBQLRAKLN--REEAHTTDLKEEDDKRMKMLEQKRG-S 273
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   192 BELPLIYNRIIGCYVTIFQNISLRDVFYREMSKLNHNLIYEMSKLEKHNSKVFVVVKGL 251
    ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   274 DSKPSLSLPR-----KTKDRRLYSISVGT 297
    ::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   252 SSSSRRSIVIPPVRTATVS----PLTGPTSPSTLSLKSESVSATEDIAPDAQGSED 307
    ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   298 EGTVTRSACQTLDTVENADMKKKLPLTMPVPKSTGSTPLVASNAKGSVCTSATWARPGID 357
    ::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   308 -NSEIKEILL-----EEERIKEGGGEASSEER---DDPLACNCPGAQAPSPTTE 352
    ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   358 RQASVGDDLIGASVPAPPPPSANKLIEENGSTGSTDPTSSTPLFSNAAPPFAQ-TTGIA 416
    ::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   353 RAKSQEEVLPGSTTPSGGALSFGSQPSSSAATEVLRRTASEG-SEQPKRASIQRTSA 411
    ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   417 PQNSQAPPWHLSHPFCANTSLHCLNFRPQA-----ASFREFQGNANDPDONGN--TTQS 468
    ::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   412 PPSRPPPPP-----RAYASPRPSGNIPSPSPTASGGGSPTSFRASLIGTGTSAPR 459
    ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   469 PPSRDVSTSRDNLVAXQLARNVTQALSFTSPQAGAFSRGPV--FTGDVGTHPPVGR 526
    ::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   460 TGLEV-----SPNPEPEKP--VRTPBEAKENENIHQNPELCSTPLMTSOVASE 508
    ||::||::||::||::||::||::||::||::||::||::||::||::||
Db   527 TLGKTGHGARVDNRGNPPPIPKKPGUSQTSFPSP-----HPQ-----LKVITSSRASRN 574
    ||::||::||::||::||::||::||::||::||::||::||::||::||

QY   509 PGKAKMDEKGNKLIKISADSSEGGQOLQVSMVPENNNTLTAPEOPEYSTENPOL 564
    ||::||::||::||::||::||::||::||::||::||::||::||::||
Db   575 TG-----AKVDNKTVASTPSS-----LPQGNRVII--NEENLPKSSSSPOL 611
    ||::||::||::||::||::||::||::||::||::||::||::||::||

```

RESULT 14  
US-10-369-493-5013  
; Sequence 5013, Application US/10369493  
; Publication No. US20030233675A1

## GENERAL INFORMATION:

APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, Xianfeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT APPLICATION NUMBER: US/10/369,493  
 PRIOR FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US 60/360,039  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 5013  
 LENGTH: 6642  
 TYPE: PRT  
 ORGANISM: Caenorhabditis elegans  
 US-10-369-493-5013

Query Match 6.9%; Score 197.5; DB 15; Length 6642;

Best Local Similarity 22.9%; Pred. No. 0.37; Indels 29; Gaps 8;  
 Matches 78; Conservative 48; Mismatches 185;

QY 232 EVMSKLEKQHSNKFVVKGLSSSSRRSLVISPVRTATVSSPLTSPSTSLKSESES 291  
 DB 1556 EIKSPVKEKSPKVEEKPAKSPKKEKSPKSPKSPKSPKSPKSPKSPKSPK 1615  
 QY 292 VSATEDLAPDAQEDNSSEIK-----ELLSEEBIEKEGSEASSEDDPLPACNGPAQAO 346  
 DB 1616 LKSPKESPEKADKPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPK 1673  
 QY 347 PSPTTE-----RAKSQEVLPSTTPSPGGALSPSGQFSSSSATEVVLTR-----TASEG 396  
 DB 1674 SSPTKTDDEVKSPKPKESQTVEEKPAKSPKKEKSPKSPKSPKSPKSPKSPK 1733  
 QY 397 SEQPKRASIOKTSAPPKPPPPRATASPRSSGNIPSSSTAGSGSPSPRASLTGTA 456  
 DB 1734 KPSPTKGEKSPKSAAEVKSPTKKEKSPKSAEKPKSPKTK-----KESPVKXAADEVK 1789  
 QY 457 SPRTSLEVP-----NPEPPKPVRTPEAKENIHONPEELCTSPTLTMSQVASEPGE 511  
 DB 1790 SP-TKEKSPKVEKPAKSPKKEKTPKSAEELKSPKKE--KSPSSPTKKTGDESKE 1846  
 QY 512 AKMEDKEKONKLIADSEGGDQLQVSMVPPNNNLTAPE 551  
 DB 1847 KSPKEPKPKSPKPKKPPGSPKPKKSPKPKKSPKPKKSPKPKKSPKPKK 1886

## RESULT 15

US-09-871-916-2  
 Sequence 2, Application US/09871916  
 Patent No. US20020019519A1

## GENERAL INFORMATION:

APPLICANT: BINGHAM, SHARON  
 APPLICANT: CASE, PATRICK  
 APPLICANT: LAWSON, SALLY NEALE  
 APPLICANT: NEWTON, RICHARD ANTHONY  
 APPLICANT: RAUSCH, OLIVER LARS  
 APPLICANT: REITH, ALASTAIR DAVID  
 APPLICANT: SANGER, GARETH JOHN  
 TITLE OF INVENTION: NEW USE  
 FILE REFERENCE: P32261-D1  
 CURRENT APPLICATION NUMBER: US/09/871,916  
 CURRENT FILING DATE: 2001-06-01  
 PRIOR APPLICATION NUMBER: 09/393,569  
 PRIOR FILING DATE: 1999-09-10  
 PRIOR APPLICATION NUMBER: GB 9907261.3  
 PRIOR FILING DATE: 1999-03-29  
 PRIOR APPLICATION NUMBER: GB 9819779.1  
 PRIOR FILING DATE: 1998-09-10  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2  
 LENGTH: 1360  
 TYPE: PRT  
 ORGANISM: HOMO SAPIENS  
 US-09-871-916-2

Query Match 6.9%; Score 197; DB 9; Length 1360;

Best Local Similarity 18.5%; Pred. No. 0.059;  
 Matches 119; Conservative 110; Mismatches 248; Indels 166; Gaps 23;

QY 15 KQVKKFSRAQKVKLOKKAIVETKDERFEQASNFYQQQAEGHK-----LYK 62  
 DB 296 RQVRIQLKDHIDRTKKGKGEKDETE---YEYSGSEEEENDSGEPSSILNLPSESTLRR 352  
 QY 63 DLKNFLSAVKVWHSESKRVSETLQBIYSSEWPGHEELKAIYVNNNDLLWEDYEELAQAV 122  
 DB 353 DFLRLQANKERSEALER-----QLEQQQORENEEHKQLL-----AEQKRIEQQE 400  
 QY 123 RTWEIYVAFSEIKERIAKGRKLVYDSARHLEAVQNAKDEAKTAKAEFF-----N 177  
 DB 401 QRRL-----EQQRREKELRKQEREQRRYEE--QVRREERERRAHEQEYIRROLE 452  
 QY 178 KACTVPEDLNQELLELPILYNSRIGCYVTFQNTSNLRDV--FYREMSKLNHLIYVMS 235  
 DB 453 EGRQLEILQQLLHQAILLLEYK-----RKQLEEQQAERLQRLKQERDYLVSQ 505  
 QY 236 KLEKQH-----SNKVYVYKGLSSSSRRSLVISP----- 263  
 DB 506 QRQEQRPVEKKPLYHYKEGSPSEKPAWAKEVEERSRLNRQSSPAMPHKVNRISSDPLNLP 565  
 QY 264 -----PVRTATVSSPLTSPSTSLKSESESVSATEDLAPDAQEDNSSEI 311  
 DB 566 PRSEFPIGQVQPARTPMDLRPV-DQIPLHVAVKSQGPALTASQSVHEQPKGLSGQE 624  
 QY 312 KELLEEBIEKEGSEASSEEDDPLPA-----CNGPAQAQ--PSPTTERAKSQEEVLP 362  
 DB 625 ALNVTSHRVMPCNSDPTSENPLTRIEKTRSSWLRQEDDIPKVPQRTTISPALA 684  
 QY 363 SSTTPSGGALSP--SQPSSSATEVVLTRTASGESEQPKKEASTORTSAPPSPPPPR 420  
 DB 685 RKNSPGNGSALGPLGSGQPIRASNPDLRT-----EPILSPLORTSSGSS----- 730  
 QY 421 ATASPRSSGNIPSS-PTASGGSGSPSPRASLTGTGTASPTSLVSP-----NPEPPEKPV 475  
 DB 731 -----SSSTPSSQFSSQGGQPGSQAGS-----SERTRVANSKSEGSFVLPHPEA 777  
 QY 476 RTPPEAKENENIHONPE-----ELCTSPILMTSQAASEGEAK 513  
 DB 778 KVKPEESRDITRPSRPASYKKAIDEDLTALAKELRELRIEETNRPMKVTDYSSSESE 837  
 QY 514 KMEDKEKONKLIADSEGGD-QLQVSMVPPNNNLTAPEPQEE 555  
 DB 838 SREEEEDGE-----SETHDGTAVVSDIPRLIPTGAPGSGNEQ 874

Search completed: March 4, 2004, 17:41:16  
 Job time : 45.1083 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:39:06 ; Search time 11.3226 Seconds

(without alignments)  
110.442 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_23\_35

Perfect score: 61

Sequence: 1 RAQEKVLQKLGKA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

1: PIR.78.\*

2: PIR1.\*

3: PIR2.\*

4: PIR3.\*

5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	61	100.0	482	2 JC5593	amphiphysin II2 -
2	61	100.0	682	1 S22700	amphiphysin - chic
3	61	100.0	695	2 S62400	amphiphysin (clone
4	41	67.2	570	1 C69985	probable DNA-depen
5	40	65.6	176	2 B75494	conserved hypotet
6	40	65.6	446	2 F83147	conserved hypotet
7	39	63.9	505	2 B93196	PTS system, IIBC c
8	39	63.9	510	2 G98062	phosphotransferase
9	39	63.9	516	2 T10000	cytochrome P450 (C
10	39	63.9	524	2 T09999	cytochrome P450 -
11	39	63.9	524	2 T09944	probable cytochrom
12	38	62.3	405	2 S68254	arrestin isoform 2
13	38	62.3	407	2 S68253	arrestin isoform 1
14	38	62.3	415	2 S68255	arrestin isoform 1
15	38	62.3	420	2 A47140	TGF beta homolog d
16	38	62.3	427	2 A40735	TGF beta homolog d
17	38	62.3	439	2 C84981	hypothetical prote
18	38	62.3	461	2 T22946	hypothetical prote
19	38	62.3	474	1 D64140	conserved hypotet
20	38	62.3	474	1 C64801	ylea protein - Esc
21	38	62.3	474	2 E82261	conserved hypotet
22	38	62.3	474	2 F85566	hypothetical prote
23	38	62.3	474	2 C90716	ylea protein - Esc
24	38	62.3	488	2 AC0584	MiaB protein (prob
25	38	62.3	497	2 F82747	conserved hypotet
26	37	60.7	60	2 B61819	conserved hypotet
27	37	60.7	272	2 H70712	hypothetical prote
28	37	60.7	375	1 G64436	N-acetyl-gamma-glu
29	37	60.7	403	2 B69196	conserved hypotet

30	37	60.7	418	2 T23861	kynurenine-oxoglut
31	37	60.7	472	2 B90095	hypothetical prote
32	37	60.7	526	1 G71081	probable helicase
33	37	60.7	1208	2 T39068	coiled coil protei
34	37	60.7	1274	2 D84485	probable retroelem
35	37	60.7	1468	1 S30818	hypothetical prote
36	36	59.0	96	2 G69037	phosphoribosyl-ATP
37	36	59.0	285	2 I38248	steroidogenic acut
38	36	59.0	285	2 J04315	steroidogenic acut
39	36	59.0	296	2 F87411	hypothetical prote
40	36	59.0	381	2 B56607	arrestin homolog -
41	36	59.0	418	2 A34851	beta-arrestin, bra
42	36	59.0	418	2 B46682	beta-arrestin 1, s
43	36	59.0	418	2 B43404	beta-arrestin1 - r
44	36	59.0	442	2 C81978	hypothetical prote
45	36	59.0	442	2 C81034	conserved hypotet
46	36	59.0	462	2 C93770	hypothetical prote
47	36	59.0	500	2 T26078	hypothetical prote
48	36	59.0	856	2 A14110	Mg2+ transport ATP
49	36	59.0	933	2 T28995	hypothetical prote
50	36	59.0	1787	2 T20160	hypothetical prote
51	35	57.4	183	2 G70450	hypothetical prote
52	35	57.4	253	2 H31006	probable oxidoredu
53	35	57.4	253	2 A85851	probable oxidoredu
54	35	57.4	294	2 F87714	chromosome partiti
55	35	57.4	313	2 S61990	hypothetical prote
56	35	57.4	339	2 E86761	conserved hypotet
57	35	57.4	341	2 F82712	integral membrane
58	35	57.4	372	2 E81350	probable RNA nucle
59	35	57.4	375	2 A39777	41K blood stage an
60	35	57.4	406	2 E72366	conserved hypotet
61	35	57.4	409	2 S18984	arrestin - human (
62	35	57.4	410	2 A59279	beta-arrestin 2 -
63	35	57.4	413	2 G86181	hypothetical prote
64	35	57.4	433	2 A71912	glycerol-3-phospha
65	35	57.4	446	1 S57900	conserved hypotet
66	35	57.4	451	2 D82702	conserved hypotet
67	35	57.4	454	2 A81244	acetyl-CoA carboxy
68	35	57.4	454	2 A11606	acetyl-CoA carboxy
69	35	57.4	504	2 T35817	probable Glu-tRNA(
70	35	57.4	515	2 A81025	ATP synthase F1, a
71	35	57.4	515	2 A81970	H+-transporting tw
72	35	57.4	522	2 B71807	DNA repair protein
73	35	57.4	570	2 A75007	DNA mismatch recog
74	35	57.4	576	2 T22700	hypothetical prote
75	35	57.4	576	2 D95207	oligonucleotidase
76	35	57.4	598	2 D98072	oligonucleotidase
77	35	57.4	695	2 T40168	hypothetical prote
78	34	55.7	227	2 S05585	tropomyosin - huma
79	34	55.7	101	2 H97263	PTS system IIB com
80	34	55.7	145	2 A81935	hypothetical prote
81	34	55.7	164	2 T24200	hypothetical prote
82	34	55.7	173	2 D82771	hypothetical prote
83	34	55.7	239	2 T51686	probable transcrip
84	34	55.7	253	1 H64981	probable dehydroge
85	34	55.7	279	2 B87674	oxidoreductase, al
86	34	55.7	296	2 T06025	transcription fact
87	34	55.7	301	2 G71929	GTP-binding protei
88	34	55.7	302	2 E84584	GTP-binding protei
89	34	55.7	313	2 T00743	glutamyl tRNA synt
90	34	55.7	329	2 I64160	hypothetical prote
91	34	55.7	332	2 AC2282	group 2 sigma 70-t
92	34	55.7	341	2 H90267	arginine deiminase
93	34	55.7	361	2 A55081	arrestin 1 - blueb
94	34	55.7	363	2 T31460	probable magnesium
95	34	55.7	364	2 A34867	arrestin - fruit f
96	34	55.7	396	2 D82171	cystathionine beta
97	34	55.7	401	2 H72765	probable flap endo
98	34	55.7	401	2 A34856	49K photoreceptor
99	34	55.7	406	2 A80111	probable monooxyge
100	34	55.7	406	2 B81331	zinc proteinase-li
101	34	55.7	423	2 H86838	serine-tRNA ligase
102	34	55.7	455	2 G70136	heat shock protein

103 34 55.7 474 2 AH0319  
104 34 55.7 535 2 S40461  
105 34 55.7 535 2 S40462  
106 34 55.7 542 1 A70220  
107 34 55.7 545 2 T10662  
108 34 55.7 585 2 T10662  
109 34 55.7 609 1 G69843  
110 34 55.7 628 2 T44581  
111 34 55.7 657 2 G71377  
112 34 55.7 662 2 T04461  
113 34 55.7 666 2 F83340  
114 34 55.7 684 2 F85075  
115 34 55.7 775 2 E70320  
116 34 55.7 779 2 A86862  
117 34 55.7 818 2 T31464  
118 34 55.7 827 2 A90544  
119 34 55.7 855 2 E75191  
120 34 55.7 917 2 B81309  
121 34 55.7 1095 1 A31225  
122 34 55.7 1331 1 XORTDH  
123 34 55.7 1335 1 XOMSDH  
124 34 55.7 1426 2 T00337  
125 33 54.1 77 2 B72271  
126 33 54.1 92 2 E64345  
127 33 54.1 94 2 D64446  
128 33 54.1 117 2 AB3298  
129 33 54.1 120 2 A81431  
130 33 54.1 141 2 AG2622  
131 33 54.1 141 2 G97404  
132 33 54.1 151 2 A75106  
133 33 54.1 165 2 AH1014  
134 33 54.1 172 2 JQ0391  
135 33 54.1 190 2 B72416  
136 33 54.1 256 2 G83953  
137 33 54.1 265 2 JC4509  
138 33 54.1 272 2 B71618  
139 33 54.1 277 2 C81231  
140 33 54.1 277 2 D82005  
141 33 54.1 280 2 A81746  
142 33 54.1 281 2 D71483  
143 33 54.1 284 2 A55455  
144 33 54.1 298 2 G90529  
145 33 54.1 318 1 D64417  
146 33 54.1 318 2 F84495  
147 33 54.1 318 2 T08695  
148 33 54.1 340 2 JC1357  
149 33 54.1 349 2 S03575  
150 33 54.1 380 2 C71688

## ALIGNMENTS

RESULT 1  
JC5593  
amphiphysin II2 - human  
C:Species: Homo sapiens (man)  
C>Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 09-Jun-2000  
C:Accession: JC5593  
R:Tsutsui, K.; Maeda, Y.; Tsutsui, K.; Seki, S.; Tokunaga, A.  
Biochem. Biophys. Res. Commun. 236, 178-183, 1997  
A:Title: cDNA cloning of a novel amphiphysin isoform and tissue-specific expression of it  
A:Reference number: JC5593; MUID:97366616; PMID:9223448  
A:Accession: JC5593  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-482 <TSU>  
A:Cross-references: DBJ:AF001383; NID:G2199534; PIDN:AAB61363.1; PID:G2199535  
A:Experimental source: fetal brain  
C:Comment: This protein is involved in the synaptic vesicle recycling and in the regulation of  
C:Superfamily: amphiphysin; RVS161 protein homology  
F:16-275/Domain: RVS161 protein homology <RVS>  
F:410-481/Domain: SH3 #status predicted <SH3>

Query Match 100.0%; Score 61; DB 2; Length 482;  
Best Local Similarity 100.0%; Pred. No. 0.0074;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGKA 13  
|||  
Db 24 RAQEKVLQKLGKA 36  
|||

## RESULT 2

S22700  
amphiphysin - chicken  
C:Species: Gallus gallus (chicken)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S22700  
R:Lichte, B.; Veb, R.W.; Meyer, H.E.; Killmann, M.W.  
EMBO J. 11, 2521-2530, 1992  
A:Title: Amphiphysin, a novel protein associated with synaptic vesicles.  
A:Reference number: S22700; MUID:92331604; PMID:1628617  
A:Accession: S22700  
A:Molecule type: mRNA  
A:Residues: 1-682 <LIC>  
A:Cross-references: EMBL:X60422; NID:G62842; PIDN:CAA42953.1; PID:G62843  
C:Superfamily: amphiphysin; RVS161 protein homology  
F:11-270/Domain: RVS161 protein homology <RVS>

Query Match 100.0%; Score 61; DB 1; Length 682;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGKA 13  
|||  
Db 19 RAQEKVLQKLGKA 31  
|||

## RESULT 3

S62400  
amphiphysin (clone 22-2) - human  
C:Species: Homo sapiens (man)  
C>Date: 12-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999  
C:Accession: S62400; I37166  
R:David, C.; Solimena, M.; de Camilli, P.  
FEBS Lett. 351, 73-79, 1994  
A:Title: Autoimmunity in Stiff-Man Syndrome with breast cancer is targeted to the C-term  
A:Reference number: S48686; MUID:94357284; PMID:8076697  
A:Accession: S62400  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-695 <DAV>  
A:Cross-references: EMBL:U07616; NID:G550449; PIDN:AAA21865.1; PID:G550450  
R:Yamamoto, R.; Li, X.; Winter, S.; Francke, U.; Kilmann, M.W.  
Hum. Mol. Genet. 4, 265-268, 1995  
A:Title: Primary structure of human amphiphysin, the dominant autoantigen of paraneoplas  
A:Reference number: I37166; MUID:95276740; PMID:7757077  
A:Accession: I37166  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-695 <RES>  
A:Cross-references: EMBL:X81438; NID:G662991; PIDN:CAA57197.1; PID:G662992  
C:Genetics:  
A:Gene: GDB:AMPH  
A:Cross-references: GDB:386990  
A:Map position: 7p14-7p13  
C:Superfamily: amphiphysin; RVS161 protein homology  
F:11-270/Domain: RVS161 protein homology <RVS>

Query Match 100.0%; Score 61; DB 2; Length 695;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGKA 13  
|||



Db 19 RAQEKVLQKLGKA 31

## RESULT 4

```

conserved hypothetical protein PA3980 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83147
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: F83147
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-446 <STO>
A:Cross-references: GB:AE004816; GB:AE004091; NID:g9950168; PIDN:AAG07367.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3980
C:Superfamily: conserved hypothetical protein b0835

Query Match          65.6%; Score 40; DB 2; Length 446;
Best Local Similarity 72.7%; Pred. No. 32;
Matches      8; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

QY      1 RAQEKVLQKLG 11
          :|||||
Db       54 KAQEKVFSLKG 64

RESULT 7

```

## RESULT 7

PTS system, IIBC components [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
R:Accession: B95196  
R:Retelling: H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.

A;Accession: D01987; Accession: A95000; MUID:21357209; PMID:11463916  
A;Accession: B95196  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-505 <KUR>  
A;Cross-references: GB:AE005672; PIDN:AAK75763.1; PID:g14973177; GSPDB:GNC00164; TIGR:SP4  
A;Experimental source: strain TIGR4  
C;Genetics:  
C;Gene: SPI684  
C;Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosphat

Query Match 63.9%; Score 39; DB 2; Length 505;  
Best Local Similarity 88.9%; Pred. No. 53;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KVLQKLGGKA 13  
||| |||  
Db 3 KVLQKVGGKA 11

phosphotransferase system enzyme II (EC 2.7.1.69), PTS-III [imported] - Streptococcus pyogenes  
G89062  
C:Species: Streptococcus pneumoniae  
C:Date: 23-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
C:Accession: G98062  
R:Hoskins, J.A.; Albom JR., W.; Arnold, J.; Blaszczyk, L.; Bargett, S.; DeHoff, B.S.; Eberhart-Keef, M.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McMahon, S.; Nishimura, R.; Leslang, D.J.; Lee, L.N.

Y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A:Authors: tang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A:Reference number: A97872; MUID:21429245; PMID:11544234  
 A:Accession: G98062  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-510 <KUR>  
 A:Cross-references: GB:AE007317; PIDN:AAL00332.1; PID:gl5459191; GSPDB:GN00174  
 C:Genetics: PTS-EII  
 A:Gene: Pts-EII  
 C:Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosphotransferase  
 C:Keywords: phosphotransferase

Query Match 63.9%; Score 39; DB 2; Length 510;  
 Best Local Similarity 88.9%; Pred. No. 54;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KYLQKLGKA 13  
 |||:|:|:|  
 DB 8 KYLQKVGKA 16  
 |||:|:|:|

RESULT 9  
 T10000  
 cytochrome P450 (CYP72C) - Madagascar periwinkle (fragment)  
 C:Species: Catharanthus roseus (Madagascar periwinkle)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001  
 C:Accession: T10000  
 R:Mangold, U.; Eichel, J.; Batschauer, A.; Lanz, T.; Spangenberg, G.; Werck-  
 Plant Sci. 96, 129-136, 1994  
 A:Title: Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharanth  
 A:Reference number: Z16915  
 A:Accession: T10000  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-516 <MAN>  
 A:Cross-references: EMBL:L19075; NID:9404689; PID:g404690  
 A:Experimental source: cv. cp3  
 C:Genetics:  
 A:Gene: CYP72C  
 C:Superfamily: unassigned cytochrome P450; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:318-481/Domain: cytochrome P450 homology <P45>  
 F:459/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 63.9%; Score 39; DB 2; Length 516;  
 Best Local Similarity 66.7%; Pred. No. 55;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAEKVLQKLGK 12  
 |||:|:|:|  
 DB 346 RAREVLQAFGK 357  
 |||:|:|:|

RESULT 10  
 T09999  
 Cytochrome P450 - Madagascar periwinkle  
 C:Species: Catharanthus roseus (Madagascar periwinkle)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001  
 C:Accession: T09999  
 R:Mangold, U.; Eichel, J.; Batschauer, A.; Lanz, T.; Spangenberg, G.; Werck-  
 Plant Sci. 96, 129-136, 1994  
 A:Title: Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharanth  
 A:Reference number: Z16915  
 A:Accession: T09999  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-524 <MAN>  
 A:Cross-references: EMBL:L19074; NID:9404687; PID:g404688  
 A:Experimental source: cv. cp3  
 C:Genetics:

A:Gene: CYP72B  
 A:Introns: 96/1; 170/3; 252/2; 381/3  
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:329-492/Domain: cytochrome P450 homology <P45>  
 F:470/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 63.9%; Score 39; DB 2; Length 524;  
 Best Local Similarity 66.7%; Pred. No. 55;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAEKVLQKLGK 12  
 |||:|:|:|  
 DB 357 RAREVLQAFGK 368  
 |||:|:|:|

RESULT 11  
 T09944  
 probable cytochrome P450 protein - Madagascar periwinkle  
 N:Alternate names: CYP72 protein  
 C:Species: Catharanthus roseus (Madagascar periwinkle)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001  
 C:Accession: T09944  
 R:Vetter, H.P.; Mangold, U.; Schroeder, G.; Warner, F.J.; Werck-Reichhart, D.; Schroeder  
 Plant Physiol. 100, 998-1007, 1992  
 A:Title: Molecular analysis and heterologous expression of an inducible cytochrome P-450  
 A:Reference number: Z16902  
 A:Accession: T09944  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-524 <VET>  
 A:Cross-references: EMBL:L10081; NID:gl67483; PID:gl67484  
 C:Genetics:  
 A:Gene: CYP72  
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:329-492/Domain: cytochrome P450 homology <P45>  
 F:470/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 63.9%; Score 39; DB 2; Length 524;  
 Best Local Similarity 66.7%; Pred. No. 55;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAEKVLQKLGK 12  
 |||:|:|:|  
 DB 357 RAREVLQAFGK 368  
 |||:|:|:|

RESULT 12  
 S68254  
 arrestin isoform 2S, erythrocyte - rainbow trout  
 C:Species: Oncorhynchus mykiss (rainbow trout)  
 C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
 C:Accession: S68254  
 R:Jahns, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motais, R.; Fievet, B.  
 Biochem. J. 316, 497-506, 1996  
 A:Title: Trout red blood cell arrestin (TRCarr), a novel member of the arrestin family:  
 A:Reference number: S68253; MUID:96257743; PMID:8687393  
 A:Accession: S68254  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-405 <JAH>  
 A:Cross-references: GB:U48410  
 C:Superfamily: arrestin

Query Match 62.3%; Score 38; DB 2; Length 405;  
 Best Local Similarity 58.3%; Pred. No. 64;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAEKVLQKLGK 12  
 |||:|:|:|  
 DB 100 RQERLUKLGQ 111  
 |||:|:|:|

C;Comment: This protein plays a role in the regulation of G-protein-coupled receptors.  
 C;Superfamily: arrestin  
 F;1-420/Product; arrestin arr3L #status predicted <MAT>  
 F;1-362,374-420/Product; arrestin arr3S #status predicted <MA2>

Query Match 62.3%; Score 38; DB 2; Length 407;  
 Best Local Similarity 58.3%; Pred. No. 65;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGK 12  
 Db 100 RQERLLKLGQ 111

RESULT 13  
 S68253  
 arrestin isoform 1S, erythrocyte - rainbow trout  
 C;Species: Oncorhynchus mykiss (rainbow trout)  
 C;Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 29-Sep-1999  
 C;Accession: S68253  
 R;Jahns, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motaïs, R.; Fievet, B.  
 Biochem. J. 316, 497-506, 1996  
 A;Title: Trout red blood cell arrestin (TRCGarr), a novel member of the arrestin family:  
 A;Reference number: S68253; MUID:96257743; PMID:8687393  
 A;Accession: S68253  
 A;Status: preliminary; nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-407 <JAH>  
 A;Cross-references: GB:U48410; NID:G1215723; PIDN:AAB16954.1; PID:G1215724  
 C;Superfamily: arrestin

Query Match 62.3%; Score 38; DB 2; Length 407;  
 Best Local Similarity 58.3%; Pred. No. 65;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGK 12  
 Db 100 RQERLLKLGQ 111

RESULT 14  
 S68255  
 arrestin isoform 1L, erythrocyte - rainbow trout  
 C;Species: Oncorhynchus mykiss (rainbow trout)  
 C;Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
 C;Accession: S68255  
 R;Jahns, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motaïs, R.; Fievet, B.  
 Biochem. J. 316, 497-506, 1996  
 A;Title: Trout red blood cell arrestin (TRCGarr), a novel member of the arrestin family:  
 A;Reference number: S68253; MUID:96257743; PMID:8687393  
 A;Accession: S68255  
 A;Status: preliminary; nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-415 <JAH>  
 A;Cross-references: GB:U48410  
 C;Superfamily: arrestin

Query Match 62.3%; Score 38; DB 2; Length 415;  
 Best Local Similarity 58.3%; Pred. No. 66;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGK 12  
 Db 100 RQERLLKLGQ 111

RESULT 15  
 A47140  
 arrestin arr3L - bovine  
 N;Contains: arrestin arr3S  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 16-Feb-1994 #sequence\_revision 27-Jun-1994 #text\_change 13-Sep-1998  
 C;Accession: A47140; B47140; JG2051  
 R;Sterne-Marr, R.; Gurevich, V.V.; Goldsmith, P.; Bodine, R.C.; Sanders, C.; Donoso, L.A.  
 J. Biol. Chem. 268, 15640-15648, 1993  
 A;Title: Polypeptide variants of beta-arrestin and arrestin3.  
 A;Reference number: A47140; MUID:93340166; PMID:8340388  
 A;Accession: A47140  
 A;Molecule type: mRNA  
 A;Residues: 1-420 <STE>  
 A;Cross-references: GB:L14641  
 A;Experimental source: brain  
 A;Note: sequence extracted from NCBI backbone (NCBIN:136006, NCBIP:136007)  
 A;Accession: B47140  
 A;Molecule type: mRNA  
 A;Residues: 1-362,374-420 <ST2>

Query Match 62.3%; Score 38; DB 2; Length 420;  
 Best Local Similarity 58.3%; Pred. No. 67;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGK 12  
 Db 100 RQERLLKLGQ 111

RESULT 16  
 A40735  
 TGP beta homolog dsl-1 - chicken  
 C;Species: Gallus gallus (chicken)  
 C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
 C;Accession: A40735  
 R;Basler, K.; Edlund, T.; Jessell, T.M.; Yamada, T.  
 Cell 73, 687-702, 1993  
 A;Title: Control of cell pattern in the neural tube: regulation of cell differentiation  
 A;Reference number: A40735; MUID:93272310; PMID:7916656  
 A;Accession: A40735  
 A;Status: preliminary  
 A;Molecule type: nucleic acid  
 A;Residues: 1-427 <BAS>  
 A;Cross-references: NID:G304379; PIDN:AAA48752.1; PID:G304380  
 A;Experimental source: spinal cord  
 A;Note: sequence extracted from NCBI backbone (NCBIN:132680, NCBIP:132681)  
 C;Superfamily: inhibin

Query Match 62.3%; Score 38; DB 2; Length 427;  
 Best Local Similarity 80.0%; Pred. No. 68;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 3 QEKVLQKLGK 12  
 Db 282 QESVLRKLGK 291

RESULT 17  
 C84981  
 hypothetical protein [imported] - Buchnera sp. (strain APS)  
 C;Species: Buchnera sp.  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C;Accession: C84981  
 R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
 Nature 407, 81-86, 2000  
 A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A  
 A;Reference number: A84930; MUID:20445173; PMID:10993077  
 A;Accession: C84981  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-439 <STO>  
 A;Cross-references: GB:AP000398; GSPDB:GN00144  
 A;Experimental source: strain APS  
 C;Genetics:  
 A;Gene: yleA; BU441  
 C;Superfamily: conserved hypothetical protein b0835

Query Match 62.3%; Score 38; DB 2; Length 439;  
 Best Local Similarity 58.3%; Pred. No. 70;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 RAQEKVLQKLGK 12  
 Db 53 KAQEKVHQLGR 64

RESULT 18  
 T22946

hypothetical protein F58G6.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000  
C:Accession: T22946  
R:Li, Lloyd, C.  
Submitted to the EMBL Data Library, December 1995  
A:Reference number: Z19641  
A:Accession: T22946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-461 <WIL>  
A:Cross-references: EMBL:Z68217; PIDN:CAA92465.1; GSPDB:GN00022; CESP:F58G6.1  
A:Experimental source: clone F58G6  
C:Genetics:  
A:Gene: CESP:F58G6.1  
A:Map position: 4  
A:Introns: 19/3; 57/3; 96/3; 130/3; 157/3; 186/3; 219/3; 271/3; 302/3; 335/3; 357/3  
C:Superfamily: amphiphysin; RVSL161 protein homology

Query Match 62.3%; Score 38; DB 2; Length 461;  
Best Local Similarity 53.8%; Pred. No. 73;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 13  
DB 15 RTKEKLEGGK 27  
:|||||:  
:|||||:

RESULT 19  
D64140  
conserved hypothetical protein HI0019 - *Haemophilus influenzae*  
C:Species: *Haemophilus influenzae*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 28-Jul-2000  
C:Accession: D64140  
R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: D64140  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-474 <TIGR>  
A:Cross-references: GB:U32687; GB:I42023; NID:gl572955; PIDN:AAC21697.1; PID:gl572963; T  
A:Experimental source: strain Rd KW20  
C:Superfamily: conserved hypothetical protein b0835

Query Match 62.3%; Score 38; DB 1; Length 474;  
Best Local Similarity 58.3%; Pred. No. 75;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
DB 54 KAQEKVPHQLGR 65  
:|||||:  
:|||||:

RESULT 20  
C64801  
yleA protein - *Escherichia coli* (strain K-12)  
C:Species: *Escherichia coli*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Mar-2002  
C:Accession: C64801  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C  
; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: C64801  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-474 <BLAT>

A:Cross-references: GB:AE000170; GB:U00096; NID:gl786875; PIDN:AAC73762.1; PID:gl786882;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yleA  
C:Superfamily: conserved hypothetical protein b0835

Query Match 62.3%; Score 38; DB 1; Length 474;  
Best Local Similarity 58.3%; Pred. No. 75;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
DB 54 KAQEKVPHQLGR 65  
:|||||:  
:|||||:

## RESULT 21

E82261  
conserved hypothetical protein VC0962 [imported] - *Vibrio cholerae* (strain N16961 serogr  
C:Species: *Vibrio cholerae*  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: E82261  
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P  
; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: E82261  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-474 <HEI>  
A:Cross-references: GB:AE004177; GB:AE003852; NID:g9655398; PIDN:AAF94124.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0962  
A:Map position: 1  
C:Superfamily: conserved hypothetical protein b0835

Query Match 62.3%; Score 38; DB 2; Length 474;  
Best Local Similarity 58.3%; Pred. No. 75;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
DB 54 KAQEKVPHQLGR 65  
:|||||:  
:|||||:

## RESULT 22

F85566  
hypothetical protein yleA [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93  
C:Species: *Escherichia coli*  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: F85566  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: F85566  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-474 <STO>  
A:Cross-references: GB:AE005174; NID:gl2513570; PIDN:AAG54994.1; GSPDB:GN00145; UWGP:Z08  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: yleA  
C:Superfamily: conserved hypothetical protein b0835

Query Match 62.3%; Score 38; DB 2; Length 474;  
Best Local Similarity 58.3%; Pred. No. 75;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12

```

Db      54 KAQEKVPHQLGR 65
      :||||| :||:
      54 KAQEKVPHQLGR 65

RESULT 23
C90716
YleA protein - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
C:Accession: C90716
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
  Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
  DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:2115231; PMID:11258796
A:Accession: C90716
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-474 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA834122.1; PID:gl3360157; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC80699
C:Superfamily: conserved hypothetical protein b0835

  Query Match      62.3%; Score 38; DB 2; Length 474;
  Best Local Similarity 58.3%; Pred. No. 75;
  Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 RAQEKVQLQK 12
      :||||| :||:
      54 KAQEKVPHQLGR 65

Db      54 KAQEKVPHQLGR 65

RESULT 24
AC0584
MiaB protein (probable tRNA-thiotransferase (or tRNA-methylthiotransferase)) miaB [impor
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0584
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
  th, S.; Conneron, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
  S.; Moule, S.; O'Gaora, P.
  Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
  A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
  A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0584
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-488 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05141.1; PID:gl6501914; GSPDB:GN00176
C:Genetics:
A:Gene: miaB
C:Superfamily: conserved hypothetical protein b0835

  Query Match      62.3%; Score 38; DB 2; Length 488;
  Best Local Similarity 58.3%; Pred. No. 77;
  Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 RAQEKVQLQK 12
      :||||| :||:
      68 KAQEKVPHQLGR 79

Db      68 KAQEKVPHQLGR 79

RESULT 25
F82747
conserved hypothetical protein XF0906 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: F82747
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

```

```

Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82747
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-497 <SIM>
A:Cross-references: GB:AE003930; GB:AE003849; NID:g9105819; PIDN:AAF83716.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; F
  Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, E
  as-Neto, E.; Docena, C.; El-Dorri, H.; Pacincani, A.P.; Ferreira, A.J.S.
  submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
  J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
  chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
  A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
  F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;
  Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
  A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
  M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
  A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0906
C:Superfamily: conserved hypothetical protein b0835

  Query Match      62.3%; Score 38; DB 2; Length 497;
  Best Local Similarity 58.3%; Pred. No. 79;
  Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 RAQEKVQLQK 12
      :||||| :||:
      99 KAQEKVPHQLGR 110

Db      99 KAQEKVPHQLGR 110

RESULT 26
B69189
conserved hypothetical protein MTH67 / MTH82 - Methanobacterium thermoautotrophicum (str
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: B69189; E69209
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
  Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.
  ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
  J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69189
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-60 <MTH>
A:Cross-references: GB:AE000798; GB:AE000799; GB:AE000666; NID:g2621112; PIDN:AAB84588.1
A:Experimental source: strain Delta H
C:Genetics: CPY1
A:Accession: B69209
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-60 <MTH2>
A:Cross-references: GB:AE000666; NID:g2621094; PIDN:AAB84581.1; PID:g2621111
A:Experimental source: strain Delta H
C:Genetics: CPY2
A:Gene: MTH67
C:Genetics: <CPY2>
A:Gene: MTH82

  Query Match      60.7%; Score 37; DB 2; Length 60;
  Best Local Similarity 80.0%; Pred. No. 15;
  Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 QEKVQLQK 12

```

Db 2 QGKTQKLGK 11  
| | | | |  
| | | | |

RESULT 27  
H70712  
hypothetical protein rv1501 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: H70712  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtrooyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: H70712  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-273 <COL>  
A:Cross-references: GB:Z79701; GB:AL123456; NID:g3261635; PIDN:CAB02015.1; PID:e264130;  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: rv1501

Query Match 60.7%; Score 37; DB 2; Length 273;  
Best Local Similarity 50.0%; Pred. No. 65;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQKVQLQKLGK 12  
| | | | |  
| | | | |

Db 45 RVQERILTEIGK 56  
| | | | |  
| | | | |

RESULT 28  
G64436  
N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: G64436  
R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Reson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: G64436  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-375 <BUL>  
A:Cross-references: GB:U67552; GB:L77117; NID:gl591737; PIDN:AAB99099.1; PID:gl591740; T  
C:Genetics:  
A:Map position: REV1036979-1035852  
A:Start codon: TTG  
C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase  
C:Keywords: oxidoreductase

Query Match 60.7%; Score 37; DB 1; Length 375;  
Best Local Similarity 80.0%; Pred. No. 89;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EKVLQKLGKA 13  
| | | | |  
| | | | |

Db 248 EXELKGLGKA 257  
| | | | |  
| | | | |

RESULT 29  
B69196  
conserved hypothetical protein MTH72 - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 11-Jan-2000

C:Accession: B69196  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qi, D.; Spadafora, R.; Vicair, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Ji, J.; Ji, J.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: B69196  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-403 <MTH>  
A:Cross-references: GB:AE000798; GB:AE000666; NID:g2621094; PIDN:AAB84576.1; PID:g2621110  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH72  
C:Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat h  
F16-49/Domain: tetratricopeptide repeat homology <TT01>  
F50-83/Domain: tetratricopeptide repeat homology <TT02>  
F84-117/Domain: tetratricopeptide repeat homology <TT03>  
F118-151/Domain: tetratricopeptide repeat homology <TT04>  
F152-185/Domain: tetratricopeptide repeat homology <TT05>  
F186-219/Domain: tetratricopeptide repeat homology <TT06>  
F220-253/Domain: tetratricopeptide repeat homology <TT07>  
F254-287/Domain: tetratricopeptide repeat homology <TT08>  
F288-321/Domain: tetratricopeptide repeat homology <TT09>  
F322-355/Domain: tetratricopeptide repeat homology <TT10>  
F356-389/Domain: tetratricopeptide repeat homology <TT11>

Query Match 60.7%; Score 37; DB 2; Length 403;  
Best Local Similarity 80.0%; Pred. No. 96;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGK 12  
| | | | |  
| | | | |

Db 362 QGKTQKLGK 371  
| | | | |  
| | | | |

RESULT 30  
T23861  
kynurenine-oxoglutarate transaminase (EC 2.6.1.7) / glutamine-phenylpyruvate transaminase  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jun-2002  
C:Accession: T23861  
R:Hembry, C.  
Submitted to the EMBL Data Library, February 1996  
A:Reference number: Z19809  
A:Accession: T23861  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-418 <WIL>  
A:Cross-references: EMBL:Z69793; PIDN:CAA93673.1; GSPDB:GN00028; CESP:R03A10.4  
A:Experimental source: clone R03A10  
C:Genetics:  
A:Gene: CESP:R03A10.4  
A:Map position: X  
A:Introns: 38/2; 142/1; 248/1; 348/1  
C:Superfamily: aspartate transaminase  
C:Keywords: aminotransferase

Query Match 60.7%; Score 37; DB 2; Length 418;  
Best Local Similarity 63.6%; Pred. No. 99;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 AOEKVLQKLGK 12  
| | | | |  
| | | | |

Db 407 AAEILKGLGK 417  
| | | | |  
| | | | |

RESULT 31  
B90095  
hypothetical protein orf472 [imported] - Guillardia theta nucleomorph  
C:Species: nucleomorph Guillardia theta  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
 C:Accession: B90095  
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Red  
 Nature 410, 1091-1096, 2001  
 A:Title: The highly reduced genome of an enslaved algal nucleus.  
 A:Reference number: A99082; MUID:11323671; PMID:11323671  
 A:Accession: B90095  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-472 <DOU>  
 A:Cross-references: GB:AF165818; NID:913794515; PIDN:AAK39890.1; GSPDB:GN00150  
 C:Genetics:  
 A:Gene: orf472  
 A:Map position: 1  
 A:Genome: nucleomorph  
 C:Keywords: nucleomorph

Query Match 60.7%; Score 37; DB 2; Length 472;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLG 11  
 DB 461 EKKVLQKIG 469

RESULT 32  
 G71081  
 Probable helicase protein PH0917 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
 C:Accession: G71081  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kuehida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: G71081  
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-526 <KAW>  
 A:Cross-references: GB:AP000004; NID:G3236131; PIDN:BAA30013.1; PID:G3257330  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0917  
 C:Superfamily: Pyrococcus horikoshii probable helicase PH0917  
 C:Keywords: ATP; nucleotide binding; P-loop  
 P:54-61/Region: nucleotide-binding motif A (P-loop)  
 P:165-170/Region: nucleotide-binding motif B  
 P:169-172/Region: DEAH motif

Query Match 60.7%; Score 37; DB 1; Length 526;  
 Best Local Similarity 69.2%; Pred. No. 1.2e+02;  
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 DB 11 RAFEDALQKLAKA 23

RESULT 33  
 T39068  
 Coiled coil protein - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T39068  
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Connor, R.E.; Wood, V.  
 submitted to the EMBL Data Library, August 1997  
 A:Reference number: 221825  
 A:Accession: T39068  
 A:Status: Preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA

A:Residues: 1-1208 <MUR>  
 A:Cross-references: EMBL:Z81317; PIDN:CA03608.1; GSPDB:GN00066; SPDB:SPACG9.06c  
 A:Experimental source: strain 972h-; cosmid c6G9  
 C:Genetics:  
 A:Gene: SPDB:SPACG9.06c  
 A:Map position: 1

Query Match 60.7%; Score 37; DB 2; Length 1208;  
 Best Local Similarity 58.3%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 DB 152 REQKVLKVR 163

RESULT 34  
 D84485  
 Probable retroelement pol polyprotein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: D84485  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84485  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1274 <STO>  
 A:Cross-references: GB:AE002093; NID:G4309763; PIDN:AAD15532.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: Atg07400  
 A:Map position: 2

Query Match 60.7%; Score 37; DB 2; Length 1274;  
 Best Local Similarity 70.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 EKVQLKLGKA 13  
 DB 1233 EKVLQVQKA 1242

RESULT 35  
 S30818  
 Hypothetical protein YER164w - yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Dec-2002  
 C:Accession: S30818; S50667  
 R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor,  
 submitted to the EMBL Data Library, February 1993  
 A:Reference number: S30812  
 A:Accession: S30818  
 A:Molecule type: DNA  
 A:Residues: 1-1468 <MUL>  
 A:Cross-references: GB:U18917; EMBL:L10718; NID:G603377; PIDN:AAB64691.1; PID:G603404  
 R:Dietrich, F.S.  
 submitted to the EMBL Data Library, December 1994  
 A:Description: The sequence of S. cerevisiae cosmid 8229, 9115, 9132, 9981, and lambda  
 A:Reference number: S50667  
 A:Accession: S50667  
 A:Molecule type: DNA  
 A:Residues: 1-1468 <DIE>  
 A:Cross-references: EMBL:U18917; NID:G603377; PIDN:AAB64691.1; PID:G603404; GSPDB:GN0000  
 C:Genetics:  
 A:Gene: SGD:CHD1; MIPS:YER164w  
 A:Cross-references: SGD:S0000966; MIPS:YER164w  
 A:Map position: 5R  
 C:Superfamily: chromodomain helicase CHD1; chromobox homology  
 C:Keywords: DNA binding; nucleus

F:195-233/Domain: chromobox homology <CB1>  
F:285-327/Domain: chromobox homology <CB2>

Query Match 60.7%; Score 37; DB 1; Length 1468;  
Best Local Similarity 70.0%; Pred. NO. 3.4e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEVLQKLG 11  
|:|:|:|:|:|:|  
DB 668 AEERVLQKFG 677

## RESULT 36

GC9037  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Jun-2002  
C:Accession: G69037  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.; Xi, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional  
A:Reference number: AG9000; MUID:96037514; PMID:9371463  
A:Accession: G69037  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-96 <MTH>  
A:Cross-references: GB:AB000893; GB:AB000666; NID:g2622375; PIDN:AA885765.1; PID:g262239  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1283  
C:Superfamily: phosphoribosyl-ATP pyrophosphatase h1se  
C:Keywords: hydrolase  
F:8-80/Region: h1se protein similarity

Query Match 59.0%; Score 36; DB 2; Length 96;  
Best Local Similarity 45.5%; Pred. NO. 35;  
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 2 AQEVLQKLG 12  
|:|:|:|:|:|:|  
DB 37 AEDKLEKIGE 47

## RESULT 37

I38248  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: I38248; I38896  
R:Sugawara, T.; Lin, D.; Holt, J.A.; Martin, K.O.; Javitt, N.B.; Miller, W.L.; Strauss, B.  
Biochemistry 34, 12506-12512, 1995  
A:Title: Structure of the human steroidogenic acute regulatory protein (STAR) gene: STAR  
A:Reference number: I38248; MUID:96038208; PMID:7547998  
A:Accession: I38248  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-285 <RES>  
A:Cross-references: EMBL:U29105; NID:gl041696; PIDN:AAC50234.1; PID:g1041698  
R:Sugawara, T.; Holt, J.A.; Driscoll, D.; Strauss III, J.F.; Lin, D.; Miller, W.L.; Patb  
Proc. Natl. Acad. Sci. U.S.A. 92, 4778-4782, 1995  
A:Title: Human steroidogenic acute regulatory protein: functional activity in COS-1 cell  
A:Reference number: I38896; MUID:95281540; PMID:7761400  
A:Accession: I38896  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-285 <RES>  
A:Cross-references: EMBL:U17280; NID:g727252; PIDN:AAC50141.1; PID:g727253  
C:Genetics:  
A:Gene: STAR  
A:Cross-references: GDB:STAR; GDB:635457; OMIM:600617  
A:Map position: sp11.2-sp11.2

A:Introns: 22/1; 60/1; 102/3; 155/3; 217/2; 248/3

Query Match 59.0%; Score 36; DB 2; Length 285;  
Best Local Similarity 87.5%; Pred. NO. 1e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLGK 12  
|:|:|:|:|:|:|  
DB 155 KVLQKLGK 162

## RESULT 38

JC4315  
steroidogenic acute regulatory protein - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 29-Nov-1995 #sequence\_revision 08-Feb-1996 #text\_change 18-Jul-1997  
C:Accession: JC4315  
R:Hartung, S.; Rust, W.; Balvers, M.; Iwell, R.  
Biochem. Biophys. Res. Commun. 215, 646-653, 1995  
A:Title: Molecular cloning and in vivo expression of the bovine steroidogenic acute reg  
A:Reference number: JC4315; MUID:96011827; PMID:7488004  
A:Accession: JC4315  
A:Molecule type: mRNA  
A:Residues: 1-285 <HAR>  
C:Comment: This protein is an acute controller of the rate-limiting transfer of cholesterol  
C:Genetics:  
A:Gene: STAR  
F:226-264/Region: metalloproteinase-1 tissue inhibitor similarity

Query Match 59.0%; Score 36; DB 2; Length 285;  
Best Local Similarity 87.5%; Pred. NO. 1e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KVLQKLGK 12  
|:|:|:|:|:|:|  
DB 155 KVLQKLGK 162

## RESULT 39

P87411  
hypothetical protein CCI309 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: P87411  
R:Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: P87411  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-296 <STO>  
A:Cross-references: GB:AE005673; NID:gl3422650; PIDN:AAK23290.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CCI309

Query Match 59.0%; Score 36; DB 2; Length 296;  
Best Local Similarity 58.3%; Pred. NO. 1.1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQKVLQKLGK 12  
|:|:|:|:|:|:|  
DB 255 RGHAKLEKLGK 266

## RESULT 40

B56607  
arrestin homolog - tobacco budworm  
C:Species: Heliothis virescens (tobacco budworm)  
C:Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 13-Sep-1998  
C:Accession: B56607



R:Raming, K.; Freitag, J.; Krieger, J.; Breer, H.

Cell. Signal. 5, 69-80, 1993

A:Title: Arrestin-subtypes in insect antennae.

A:Reference number: A56607; MUID:9319955; PMID:8452755

A:Accession: B56607

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-381 <RAM>

A:Experimental source: antennae

A>Note: sequence extracted from NCBI backbone (NCBIP:127926)

C:Superfamily: arrestin

Query Match 59.0%; Score 36; DB 2; Length 381;

Best Local Similarity 54.5%; Pred. No. 1.4e+02;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLG 11

DB 98 RTQERLIKKLG 108

RESULT 41

A34851

Beta-arrestin, brain - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 13-Jul-1990 #sequence\_revision 21-Oct-1992 #text\_change 29-Sep-1999

C:Accession: A34851; JC2052

R:Lohe, M.J.; Benovic, J.L.; Codina, J.; Caron, M.G.; Lefkowitz, R.J.

Science 248, 1547-1550, 1990

A:Title: Beta-arrestin: a protein that regulates beta-adrenergic receptor function.

A:Reference number: A34851; MUID:90296080; PMID:2163110

A:Accession: A34851

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-418 <LOH>

A:Cross-references: GB:M33601; NID:g162691; PIDN:AAA30387.1; PID:g162692

A:Experimental source: brain

R:Stierne-Marr, R.; Gurevich, V.V.; Goldsmith, P.; Bodine, R.C.; Sanders, C.; Donoso, L.A.

J. Biol. Chem. 268, 15640-15648, 1993

A:Title: Polypeptide variants of beta-arrestin and arrestin3.

A:Reference number: A47140; MUID:93340166; PMID:8340388

A:Accession: JC2052

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-15, 'C', 17-30, 'L', 32, 'K', 34-83, 'Y', 85-176, 'K', 178-206, 'L', 208-333, 342-418 <

A:Experimental source: lung, spleen

C:Comment: This protein is an arrestin homolog which acts analogously to desensitize the

C:Superfamily: arrestin

F:1-418/Product: beta-arrestin, brain #status predicted <MAT>

F:1-333, 342-418/Product: beta-arrestin, lung #status predicted <MA2>

Query Match

59.0%; Score 36; DB 2; Length 418;

Best Local Similarity 50.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLG 12

DB 99 RLQERLIKKLG 110

RESULT 42

B46682

Beta-arrestin 1, splice form A - human

N:Contains: beta-arrestin 1, splice form B

C:Species: Homo sapiens (man)

C>Date: 22-Sep-1993 #sequence\_revision 22-Sep-1993 #text\_change 02-Jun-2000

C:Accession: B46682; A46682

R:Parutti, G.; Peracchia, F.; Salles, M.; Ambrosini, G.; Masini, M.; Rotilio, D.; De Bl

J. Biol. Chem. 268, 9753-9761, 1993

A:Title: Molecular analysis of human beta-arrestin-1: cloning, tissue distribution, and

A:Reference number: A46682; MUID:93252853; PMID:846659

A:Accession: B46682

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-418 <PAR1>

A:Cross-references: GB:L04685; NID:g347956; PIDN:AAA35559.1; PID:g347958

A>Note: sequence extracted from NCBI backbone (NCBIP:134014)

A:Accession: A46682

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-333, 342-418 <PAR2>

A:Cross-references: GB:L04685; NID:g347956; PIDN:AAA35559.1; PID:g347957

A>Note: sequence extracted from NCBI backbone (NCBIP:134015)

C:Genetics:

A:Gene: GDB:ARRB1; ARRB1

A:Cross-references: GDB:L141541; OMIM:107940

A:Map position: 11q13-11q13

C:Superfamily: arrestin

C:Keywords: alternative splicing

Query Match 59.0%; Score 36; DB 2; Length 418;

Best Local Similarity 50.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLG 12

DB 99 RLQERLIKKLG 110

RESULT 43

B43404

Beta-arrestin1 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 29-Sep-1999

C:Accession: B43404

R:Attramadal, H.; Arriza, J.L.; Aoki, C.; Dawson, T.M.; Codina, J.; Kwatra, M.M.; Snyder

J. Biol. Chem. 267, 17882-17890, 1992

A:Title: Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene family.

A:Reference number: A43404; MUID:92388146; PMID:1517224

A:Accession: B43404

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-418 <ATT>

A:Cross-references: GB:M91589; NID:g949985; PIDN:AAA74459.1; PID:g203102

A>Note: sequence extracted from NCBI backbone (NCBIP:112791)

C:Superfamily: arrestin

Query Match 59.0%; Score 36; DB 2; Length 418;

Best Local Similarity 50.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVLQKLG 12

DB 99 RLQERLIKKLG 110

RESULT 44

C81978

hypothetical protein NMA0590 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001

C:Accession: C81978

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:20222536; PMID:10761919

A:Accession: C81978

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-442 <PAR>

A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83881.1; PID:g737932

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA0590

C:Superfamily: conserved hypothetical protein b0835

Query Match 59.0%; Score 36; DB 2; Length 442;  
 Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 12  
 :|||||:  
 Db 54 KAQKVFSLGR 65

## RESULT 45

C81034  
 conserved hypothetical protein NMB1866 [imported] - Neisseria meningitidis (strain MC58)  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: C81034  
 R:Telletlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: AB1000; MUID:20175755; PMID:10710307  
 A:Accession: C81034  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-442 <TET>  
 A:Cross-references: GB:AE002536; GB:AE002098; NID:G7227115; PIDN:AAF42200.1; PID:G722712  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1866  
 C:Superfamily: conserved hypothetical protein b0835

Query Match 59.0%; Score 36; DB 2; Length 442;  
 Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 12  
 :|||||:  
 Db 54 KAQKVFSLGR 65

## RESULT 46

C89770  
 hypothetical protein SA0099 [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 06-Jan-2003  
 C:Accession: C89770  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
 ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: AB9758; MUID:21311952; PMID:11418146  
 A:Accession: C89770  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-462 <KUR>  
 A:Cross-references: GB:BA000018; PID:G13700019; PIDN:BA41318.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: SA0099  
 C:Superfamily: multidrug-efflux transporter

Query Match 59.0%; Score 36; DB 2; Length 462;  
 Best Local Similarity 63.6%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 11  
 :|||||:  
 Db 316 RVGEKLLQKMG 326

## RESULT 47

T26078  
 hypothetical protein W02A2.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T26078  
 R:Ainscough, R.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z20148  
 A:Accession: T26078  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-500 <WIL>  
 A:Cross-references: EMBL:Z82286; PIDN:CAB05307.1; GSPDB:GN00022; CESP:W02A2.4  
 A:Experimental source: clone W02A2  
 C:Genetics:  
 A:Gene: CESP:W02A2.4  
 A:Map position: 4  
 A:Introns: 57/3; 190/3; 375/3

Query Match 59.0%; Score 36; DB 2; Length 500;  
 Best Local Similarity 53.8%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 13  
 :|||||:  
 Db 252 QATEVIVKLGKA 264

## RESULT 48

AH1410  
 Mg2+ transport ATPase homolog lmo2689 [imported] - Listeria monocytogenes (strain EGD-e)  
 C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 17-May-2002  
 C:Accession: AH1410  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Anand, A.; Baquero, F.; Berche, P.; Bloeker  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AH1410  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-856 <GLA>  
 A:Cross-references: GB:NC 003210; PIDN:CAD00902.1; PID:G16412189; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo2689  
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

Query Match 59.0%; Score 36; DB 2; Length 856;  
 Best Local Similarity 77.8%; Pred. No. 3e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 QEKVLQKLG 11  
 :|||||:  
 Db 21 KEKVLKLG 29

## RESULT 49

T28995  
 hypothetical protein ZC513.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T28995  
 R:Wu, X.; Le, T.T.  
 submitted to the EMBL Data Library, April 1996  
 A:Description: The sequence of C. elegans cosmid ZC513.  
 A:Reference number: Z20551  
 A:Accession: T28995

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-933 <WUX>  
A:Cross-references: EMBL:U53155; PIDN:AAC48266.1; GSPDB:GN00023; CESP:ZC513.4  
A:Experimental source: strain Bristol N2; clone ZC513  
C:Genetics:  
A:Gene: CESP:ZC513.4  
A:Map position: 5  
A:Introns: 384/3; 425/2; 645/3; 689/3  
C:Superfamily: valine-tRNA ligase

Query Match 59.0%; Score 36; DB 2; Length 933;  
Best Local Similarity 63.6%; Pred. No. 3.2e+02;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLQKLGK 12  
DB 322 AREKVIQKLE 332

RESULT 50  
T20160  
hypothetical protein T14G8.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T20160; T24924  
R:McMurray, A.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19231  
A:Accession: T20160  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1787 <WIL>  
A:Cross-references: EMBL:Z67881; PIDN:CAA91798.1; GSPDB:GN00028; CESP:T14G8.1  
A:Experimental source: clone C52G5  
R:Matthews, P.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19955  
A:Accession: T24924  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1787 <W12>  
A:Cross-references: EMBL:Z67884; PIDN:CAA91810.1; GSPDB:GN00028; CESP:T14G8.1  
A:Experimental source: clone T14G8  
C:Genetics:  
A:Gene: CESP:T14G8.1  
A:Map position: x  
A:Introns: 112/3; 453/3; 597/3; 815/2; 1258/3; 1682/2; 1709/3; 1764/1

Query Match 59.0%; Score 36; DB 2; Length 1787;  
Best Local Similarity 80.0%; Pred. No. 6.1e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QEKVLQKLGK 12  
DB 1226 QETELQKLGK 1235

Search completed: March 4, 2004, 17:46:55  
Job time : 16.3226 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:32:31 ; Search time 8.80645 Seconds

(without alignments)  
76.865 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_23\_35

Perfect score: 61

Sequence: 1 RAQKVLQKLGKA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	588	1 BIN1_MOUSE	O08539 mus musculus
2	61	100.0	588	1 BIN1_RAT	O08839 rattus norv
3	61	100.0	593	1 BIN1_HUMAN	O00499 homo sapien
4	61	100.0	682	1 AMPH_CHICK	P50478 gallus gall
5	61	100.0	683	1 AMPH_RAT	O08838 rattus norv
6	61	100.0	695	1 AMPH_HUMAN	P49418 homo sapien
7	40	65.6	446	1 Y310_PSEAE	O51470 pseudomonas
8	39	63.9	524	1 CP72_CATRO	O05047 catharanthu
9	38	62.3	405	1 ARR2_ONCMY	P51467 oncorhynch
10	38	62.3	407	1 ARR1_ONCMY	P51466 oncorhynch
11	38	62.3	415	1 ARR3_ONCMY	P51468 oncorhynch
12	38	62.3	420	1 ARR2_BOVIN	P32120 bos taurus
13	38	62.3	427	1 DSL1_CHICK	P34822 gallus gall
14	38	62.3	439	1 Y426_BUCAP	O8K9C2 buchnera ap
15	38	62.3	439	1 Y441_BUCAL	P57516 buchnera ap
16	38	62.3	474	1 Y401_PASMU	O31699 pasteurella
17	38	62.3	474	1 YLEA_ECOLI	P77645 escherichia
18	38	62.3	474	1 YLEA_HAEIN	P77645 escherichia
19	37	60.7	273	1 YF01_MYCTU	P71782 mycobacteri
20	37	60.7	341	1 ARGC_METUA	Q58496 methanococc
21	37	60.7	362	1 P35_MYCPE	O50367 mycoplasma
22	37	60.7	1208	1 PCPI_SCHPO	Q32351 schizosacch
23	37	60.7	1468	1 CHD1_YEAST	P32657 saccharomyc
24	36	59.0	96	1 H822_METH	O73344 methanobact
25	36	59.0	155	1 RR7_GINBI	Q59f13 ginkgo bilo
26	36	59.0	284	1 STAR_MESAU	P70114 mesocricetu
27	36	59.0	285	1 STAR_BOVIN	Q28918 bos taurus
28	36	59.0	285	1 STAR_HORSE	O46689 equus cabal
29	36	59.0	285	1 STAR_HUMAN	P49675 homo sapien
30	36	59.0	285	1 STAR_SHEEP	P79245 ovis aries
31	36	59.0	298	1 ARRA_THETN	Q8R50 thermoanaer
32	36	59.0	381	1 ARRH_HELVI	P55274 heliothis v
33	36	59.0	410	1 ARRI_RABIT	Q95223 oryctolagus

34	36	59.0	418	1 ARRI_BOVIN	P17870 bos taurus
35	36	59.0	418	1 ARRI_HUMAN	P49407 homo sapien
36	36	59.0	418	1 ARRI_RAT	P29066 rattus norv
37	36	59.0	777	1 HEXA_PORGI	P49008 porphyromon
38	36	59.0	1787	1 CHD3_CABEL	Q22516 caenorhabdi
39	35	57.4	285	1 STAR_PIG	Q28996 sus scrofa
40	35	57.4	289	1 STAR_XENLA	Q94908 xenopus lae
41	35	57.4	294	1 PARB_CAUCR	O05190 caulobacter
42	35	57.4	313	1 Y006_YEAST	Q12094 saccharomyc
43	35	57.4	409	1 ARRI_HUMAN	P32121 homo sapien
44	35	57.4	410	1 ARRI_MOUSE	Q91Y14 mus musculu
45	35	57.4	410	1 ARRI_RAT	P29067 rattus norv
46	35	57.4	459	1 ARLY_OCEIH	O8elt9 oceanobacil
47	35	57.4	504	1 GATB_STRCO	Q22578 streptomyce
48	35	57.4	522	1 RECN_HELPJ	Q92180 helicobacte
49	35	57.4	636	1 NAF1_HUMAN	Q15025 homo sapien
50	35	57.4	647	1 NAF1_MOUSE	Q9wu8 mus musculu
51	34	55.7	253	1 YOHF_ECOLI	P33368 escherichia
52	34	55.7	285	1 STAR_BRARE	Q9dgl0 brachydanio
53	34	55.7	301	1 ERA_HELPJ	Q9zlw0 helicobacte
54	34	55.7	301	1 ERA_HELPY	P56059 helicobacte
55	34	55.7	329	1 ISPB_HAEIN	P44916 haemophilus
56	34	55.7	351	1 FEN_AERPE	Q9Yfy5 aeropyrum p
57	34	55.7	363	1 ARRI_CALVI	P51486 calliphora
58	34	55.7	364	1 ARRA_DROME	P15372 drosophila
59	34	55.7	401	1 ARRB_DROME	P19107 drosophila
60	34	55.7	423	1 SYS_IACIA	Q9cex2 lactococcus
61	34	55.7	448	1 HSLU_BORBU	Q44772 borrelia bu
62	34	55.7	468	1 TIG_STRCO	Q9f314 streptomyce
63	34	55.7	469	1 LEU2_RALSO	Q8xxx3 ralstonia s
64	34	55.7	514	1 ATPA_THIFE	P41167 thiobacillu
65	34	55.7	535	1 TCE1_AVEA	P40412 avena sativ
66	34	55.7	535	1 TCE2_AVEA	P54411 avena sativ
67	34	55.7	535	1 TCPE_ARATH	O04450 arabidopsis
68	34	55.7	539	1 CH60_FUSNP	O89J00 fusbacteri
69	34	55.7	574	1 C306_DROME	Q9vvr5 drosophila
70	34	55.7	585	1 C08A_RABIT	P98136 oryctolagus
71	34	55.7	609	1 PEPF_BACSU	O31605 bacillus su
72	34	55.7	657	1 GRAD_TREPA	O83062 treponema p
73	34	55.7	917	1 SVI_CAMJE	P41257 campylobact
74	34	55.7	1095	1 RIPA_DROME	P13217 drosophila
75	34	55.7	1330	1 XDH_RAT	P22985 rattus norv
76	34	55.7	1335	1 XDH_MOUSE	Q00519 mus musculu
77	34	55.7	1756	1 PEPL_HUMAN	O60437 homo sapien
78	33	54.1	80	1 CLVA_STYCL	P80710 styela clav
79	33	54.1	92	1 Y366_METJA	O57812 methanococc
80	33	54.1	120	1 RBFA_CAMJE	Q9piz0 campylobact
81	33	54.1	138	1 ATPE_THETN	Q8rc14 thermoanaer
82	33	54.1	151	1 REG6_PTRAB	Q9vov9 pyrococcus
83	33	54.1	172	1 MBEB_ECOLI	P13659 escherichia
84	33	54.1	175	1 CYT2_STOHE	P07845 stoichactis
85	33	54.1	265	1 FLIR_TREPA	P74932 treponema p
86	33	54.1	280	1 PARB_CHLMU	O9pln9 chlamydia m
87	33	54.1	281	1 PARB_CHLTR	O84694 chlamydia t
88	33	54.1	281	1 STAR_MOUSE	P51557 mus musculu
89	33	54.1	318	1 Y940_METJA	Q58350 methanococc
90	33	54.1	349	1 RPAI_HALMO	P15349 halococcus
91	33	54.1	364	1 RSPD_LALCL	P58290 lactococcus
92	33	54.1	380	1 LPXB_RICPR	Q9zdk7 rickettsia
93	33	54.1	386	1 RPSB_LACLA	O04506 lactococcus
94	33	54.1	389	1 CC37_DROME	Q24276 drosophila
95	33	54.1	396	1 ARRS_RANCA	P51478 rana catesb
96	33	54.1	396	1 ARRS_RANPI	P51479 rana pipien
97	33	54.1	396	1 ARRS_XENLA	P51477 xenopus lae
98	33	54.1	397	1 ACK2_VIBPA	Q871j5 vibrio para
99	33	54.1	397	1 ACK2_VIBVU	O8d7k4 vibrio vuln
100	33	54.1	400	1 ARRB_CALVI	P51487 calliphora
101	33	54.1	401	1 ARRB_DROMI	P19108 drosophila
102	33	54.1	404	1 ARRS_BOVIN	P08168 bos taurus
103	33	54.1	429	1 GDF2_HUMAN	Q9uk05 homo sapien
104	33	54.1	448	1 ZFG1_CHICK	Q92010 gallus gall
105	33	54.1	449	1 ZFG1_HUMAN	O43829 homo sapien
106	33	54.1	449	1 ZFG1_MOUSE	Q08376 mus musculu



Query Match 100.0%; Score 61; DB 1; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 0.0054;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 DB 24 RAQEKVLQKLGKA 36

RESULT 2  
 BIN1\_RAT  
 ID BIN1\_RAT STANDARD; PRT; 588 AA.  
 AC O08839;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE MYC box dependent interacting protein 1 (Bridging integrator 1)  
 DE (Amphiphysin-like protein) (Amphiphysin II).  
 GN BIN1 OR AMPHL OR AMPH2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., SUBUNIT, AND ALTERNATIVE SPLICING.  
 RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex, and Kidney;  
 RX MEDLINE=98009145; PubMed=9348539;  
 RA Wigge P., Kohler K., Hunt S., Doyle C., McMahon H.T.;  
 RA McMahon H.T., Wigge P., Smith C.;  
 RT "Amphiphysin heterodimers: potential role in clathrin-mediated  
 endocytosis."  
 RL Mol. Biol. Cell 8:2003-2015(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM AMPH2-1).  
 RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex;  
 RX MEDLINE=97424383; PubMed=9280305;  
 RA McMahon H.T., Wigge P., Smith C.;  
 RT "Clathrin interacts specifically with amphiphysin and is displaced by  
 dynamin."  
 RL PNAS Lett. 413:319-322(1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 506-588, FUNCTION, AND  
 INTERACTION WITH DNMI.  
 RX MEDLINE=98409533; PubMed=9736607;  
 RA Owen D.J., Wigge P., Vallis Y., Moore J.D.A., Evans P.R.,  
 RA McMahon H.T.;  
 RT "Crystal structure of the amphiphysin-2 SH3 domain and its role in the  
 prevention of dynamin ring formation."  
 RL EMBO J. 17:5273-5285(1998).  
 CC -!- FUNCTION: May be involved in regulation of synaptic vesicle  
 endocytosis. May act as a tumor suppressor and inhibits  
 malignant cell transformation.  
 CC -!- SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).  
 CC Binds to SYN1 and DNMI through its SH3 domain, and to clathrin  
 through a region outside of the SH3 domain. Also binds to alpha-  
 adaptin. Interacts with the N-terminal transactivation domain of  
 MYC in a manner requiring the integrity of the conserved MYC box  
 regions 1 and 2.  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=6;  
 CC Name=AMPH2-1;  
 CC IsoId=O08839-1; Sequence=Displayed;  
 CC Name=AMPH2-2;  
 CC IsoId=O08839-2; Sequence=VSP\_000260;  
 CC Name=AMPH2-3;  
 CC IsoId=O08839-3; Sequence=VSP\_000258;  
 CC Name=AMPH2-4;  
 CC IsoId=O08839-4; Sequence=VSP\_000256, VSP\_000257;  
 CC Name=AMPH2-5;  
 CC IsoId=O08839-5; Sequence=VSP\_000259;  
 CC Name=AMPH2-6;  
 CC IsoId=O08839-6; Sequence=VSP\_000256, VSP\_000259;  
 CC -!- TISSUE SPECIFICITY: Isoform AMPH2-1 is expressed in brain,

CC concentrated at nerve terminals. Isoform AMPH2-2 is widely  
 CC expressed  
 CC -!- PTM: Phosphorylated by protein kinase C.  
 CC -!- SIMILARITY: Contains 1 BAR domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; Y13380; CAA73807.1; --  
 CC PDB; 1BB9; 17-JUN-98.  
 CC InterPro; IPR003005; Amphiphysin.  
 CC InterPro; IPR006632; BAR.  
 CC InterPro; IPR004148; BAR\_dom.  
 CC InterPro; IPR001452; SH3.  
 CC Pfam; PF03114; BAR; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC PRINTS; PR01251; AMPHIPHYYSIN.  
 CC PRINTS; PR00452; SH3DOMAIN.  
 CC ProDom; PD000066; SH3; 1.  
 CC SMART; SM00721; BAR; 1.  
 CC SMART; SM00326; SH3; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 CC K0 Alternative splicing; SH3 domain; Coiled coil; Endocytosis;  
 KW Anti-oncogene; Differentiation; Phosphorylation; 3D-structure.  
 FT DOMAIN 15 42  
 FT DOMAIN 193 274  
 FT DOMAIN 379 422  
 FT DOMAIN 515 588  
 FT VARSPLIC 173 205  
 FT Missing (in isoform AMPH2-4 and isoform  
 FT AMPH2-6).  
 FT /FTId=VSP\_000256.  
 FT Missing (in isoform AMPH2-4).  
 FT /FTId=VSP\_000257.  
 FT Missing (in isoform AMPH2-3).  
 FT /FTId=VSP\_000258.  
 FT Missing (in isoform AMPH2-5 and isoform  
 FT AMPH2-6).  
 FT /FTId=VSP\_000259.  
 FT Missing (in isoform AMPH2-2).  
 FT /FTId=VSP\_000260.  
 SQ SEQUENCE 588 AA; 64533 MW; 164AC90E09547F1A CRC64;  
 Query Match 100.0%; Score 61; DB 1; Length 588;  
 Best Local Similarity 100.0%; Pred. No. 0.0054;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
 DB 24 RAQEKVLQKLGKA 36

RESULT 3  
 BIN1\_HUMAN  
 ID BIN1\_HUMAN STANDARD; PRT; 593 AA.  
 AC O0499; O00297; O00545; O43867; O60552; O60553; O60554; O60555;  
 AC O75514; O75515; O75516; O75517; O75518; O92944; Q99688;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE MYC box dependent interacting protein 1 (Bridging integrator 1)  
 DE (Amphiphysin-like protein) (Amphiphysin II) (Box-dependent myc-  
 DE interacting protein-1).  
 GN BIN1 OR AMPHL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1] SEQUENCE FROM N.A. (ISOFORM IIA).  
RP TISSUE=Brain;  
RX MEDLINE=97341217; PubMed=9195986;  
RA Ramjaun A.R., Micheva K.D., Bouchelet I., McPherson P.S.;  
RT "Identification and characterization of a nerve terminal-enriched  
RT amphiphysin isoform";  
RL J. Biol. Chem. 272:16700-16706(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS IIA AND BIN1).  
RC TISSUE=Brain, and Skeletal muscle;  
RX MEDLINE=97327761; PubMed=9182667;  
RA Butler M.H., David C., Ochoa G.-C., Freyberg Z., Daniell L., Grabs D.,  
RA Cremona O., De Camilli P.;  
RT "Amphiphysin II (SH3p9; BIN1), a member of the amphiphysin/Rvs family,  
RT is concentrated in the cortical cytomatrix of axon initial segments  
RT and nodes of ranvier in brain and around T tubules in skeletal  
RT muscle";  
RL J. Cell Biol. 137:1355-1367(1997).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM BIN1).  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=96376973; PubMed=8782822;  
RA Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
RT "BIN1 is a novel Myc-interacting protein with features of a tumour  
RT suppressor";  
RL Nat. Genet. 14:69-76(1996).  
RN [4]  
RP REVISIONS TO N-TERMINUS.  
RA Sakamuro D., Elliott K.J., Wechsler-Reya R., Prendergast G.C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS IIB; IIC1; IIC2 AND IID).  
RC TISSUE=Brain;  
RX MEDLINE=98264340; PubMed=9603201;  
RA Ramjaun A.R., McPherson P.S.;  
RT "Multiple amphiphysin II splice variants display differential clathrin  
RT binding: identification of two distinct clathrin-binding sites";  
RL J. Neurochem. 70:2369-2376(1998).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORMS I12 AND I13).  
RC TISSUE=Brain;  
RX MEDLINE=97366618; PubMed=9223448;  
RA Tsutsui K., Maeda Y., Tsutsui K., Seki S., Tokunaga A.;  
RT "cDNA cloning of a novel amphiphysin isoform and tissue-specific  
RT expression of its multiple splice variants";  
RL Biochem. Biophys. Res. Commun. 236:178-183(1997).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORMS I13; I13; BIN1-10-13; BIN1-13 AND  
RP BIN1+12A).  
RC TISSUE=Fibroblast;  
RX MEDLINE=98058932; PubMed=9395479;  
RA Wechsler-Reya R.J., Sakamuro D., Zhang J., DuHadaway J.,  
RA Prendergast G.C.;  
RT "Structural analysis of the human BIN1 gene. Evidence for tissue-  
RT specific transcriptional regulation and alternate RNA splicing";  
RL J. Biol. Chem. 272:31453-31458(1997).  
RN [8]  
RP SEQUENCE FROM N.A. (ISOFORM I12).  
RA Zhang J., Du W., Wechsler-Reya R.J., DuHadaway J., Sakamuro D.,  
RA Prendergast G.C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP SEQUENCE OF 133-593 FROM N.A.  
RC TISSUE=Brain;  
RX Yu W., Gibbs R.A.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [10]  
RP CHARACTERIZATION.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=98078712; PubMed=9418903;  
RA Wechsler-Reya R.J., Elliott K.J., Prendergast G.C.;  
RT "A role for the putative tumor suppressor Bin1 in muscle cell

RT differentiation";  
RL Mol. Cell. Biol. 18:566-575(1998).  
CC [-] FUNCTION: May be involved in regulation of synaptic vesicle  
CC endocytosis. May act as a tumor suppressor and inhibits  
CC malignant cell transformation.  
CC [-] SUBUNIT: Heterodimer with AMPH. Binds SH3GLB1 (By similarity).  
CC Binds to SYNJ1 and DNMI1 through its SH3 domain, and to clathrin  
CC through a region outside of the SH3 domain. Also binds to alpha-  
CC adaptin. Interacts with the N-terminal transactivation domain of  
CC MYC in a manner requiring the integrity of the conserved MYC box  
CC regions 1 and 2.  
CC [-] SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Isoform IIA is  
CC found in the cytoplasm while isoform BIN1 is nuclear.  
CC [-] ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=11;  
CC Comment-Additional isoforms seem to exist;  
CC Names=IIA;  
CC IsoId=O00499-1; Sequence=Displayed;  
CC Name=IIB;  
CC IsoId=O00499-2; Sequence=VSP\_000246, VSP\_000252;  
CC Name=IIC1;  
CC IsoId=O00499-3; Sequence=VSP\_000249;  
CC Name=IIC2;  
CC IsoId=O00499-4; Sequence=VSP\_000246, VSP\_000249;  
CC Name=IID;  
CC IsoId=O00499-5; Sequence=VSP\_000248;  
CC Name=II2;  
CC IsoId=O00499-6; Sequence=VSP\_000246, VSP\_000253;  
CC Name=II3;  
CC IsoId=O00499-7; Sequence=VSP\_000246, VSP\_000250;  
CC Name=BIN1;  
CC IsoId=O00499-8; Sequence=VSP\_000246, VSP\_000247, VSP\_000250;  
CC Name=BIN1-10-13;  
CC IsoId=O00499-9; Sequence=VSP\_000246, VSP\_000251;  
CC Name=BIN1-13;  
CC IsoId=O00499-10; Sequence=VSP\_000246, VSP\_000247, VSP\_000251;  
CC Name=BIN1+12A;  
CC IsoId=O00499-11; Sequence=VSP\_000246, VSP\_000247, VSP\_000253;  
CC [-] TISSUE SPECIFICITY: ISOFORM IIA IS EXPRESSED ONLY IN THE BRAIN  
CC WHERE IT IS CONCENTRATED IN AXON INITIAL SEGMENTS AND NODES OF  
CC RANVIER. ISOFORM BIN1 IS WIDELY EXPRESSED WITH HIGHEST EXPRESSION  
CC IN SKELETAL MUSCLE.  
CC [-] PTM: Phosphorylated by protein kinase C (By similarity).  
CC [-] SIMILARITY: Contains 1 BAR domain.  
CC [-] SIMILARITY: Contains 1 SH3 domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF004015; AAC51345.1; -  
CC EMBL; AF070576; AAC28646.1; -  
CC EMBL; U68485; AAC17461.1; -  
CC EMBL; AF001383; AAC61363.1; -  
CC EMBL; AF043898; AAC39710.1; -  
CC EMBL; AF043899; AAC39711.1; -  
CC EMBL; AF043900; AAC39712.1; -  
CC EMBL; AF043901; AAC39713.1; -  
CC EMBL; U87558; AAB63263.1; -  
CC EMBL; U87589; AAC23440.1; -  
CC EMBL; AF068914; AAC24126.1; -  
CC EMBL; AF068915; AAC24127.1; -  
CC EMBL; AF068916; AAC24128.1; -  
CC EMBL; AF068917; AAC23750.1; -  
CC EMBL; AF068918; AAC23751.1; -  
CC EMBL; U84004; AAC23440.1; -  
CC EMBL; U83999; AAC23440.1; JOINED.  
CC EMBL; U84001; AAC23440.1; JOINED.  
CC EMBL; U84002; AAC23440.1; JOINED.  
CC EMBL; U84003; AAC23440.1; JOINED.

```

DR EMBL; U84004; AAC23441.1; ALT INIT.
DR EMBL; U83999; AAC23441.1; JOINED.
DR EMBL; U84001; AAC23441.1; JOINED.
DR EMBL; U84002; AAC23441.1; JOINED.
DR EMBL; U84003; AAC23441.1; JOINED.
DR FIR; JCS593; JCS593.
DR TRANSFAC; T03490; -.
DR Genew; HGNC:1052; BIN1.
DR MIM; 601248; -.
DR GO; GO:0015629; C-actin cytoskeleton; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR GO; GO:0006899; P:nonselective vesicle transport; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR004148; BAR_dom.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR ProDom; PD000066; SH3; 1.
DR ProSITE; PS50002; SH3; 1.
KW Alternative splicing; SH3 domain; Coiled coil; Endocytosis;
KW Anti-oncogene; Differentiation; Phosphorylation.
FT DOMAIN 15 42
FT COILED COIL (POTENTIAL).
FT DOMAIN 193 267
FT COILED COIL (POTENTIAL).
FT DOMAIN 378 421
FT CLATHRIN-BINDING.
FT DOMAIN 520 592
FT SH3.
FT VARSPLIC 174 204
FT Missing (in isoform IIB, isoform IIC2,
FT isoform I12, isoform I13, isoform BIN1,
FT isoform BIN1+12A, isoform BIN1-10-13 and
FT isoform BIN1-13).
FT P -> PRKSKLFGSLRRKK (in isoform BIN1,
FT isoform BIN1+12A and isoform BIN1-13).
FT /FTID-VSP 000246.
FT VARSPLIC 285 285
FT Missing (in isoform IIC1 and isoform
FT IIC2).
FT /FTID-VSP 000247.
FT VARSPLIC 335 377
FT Missing (in isoform IID).
FT /FTID-VSP 000248.
FT VARSPLIC 335 421
FT Missing (in isoform IIC1 and isoform
FT IIC2).

Query Match 100.0%; Score 61; DB 1; Length 593;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
Db 24 RAQEKVLQKLGKA 36

RESULT 4
AMPH_CHICK
ID AMPH CHICK STANDARD; PRT; 682 AA.
AC P50478;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amphiphysin.
GN AMPH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Tetra-hybrid; TISSUE=Forebrain;
RX MEDLINE=92331604; PubMed=1628617;
RA Lichte B., Veh R.W., Meyer H.E., Killmann M.W.;
RT "Amphiphysin, a novel protein associated with synaptic vesicles.";
RL EMBL J. 11:2521-2530(1992).
CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in
synapses and certain endocrine cell types. May control the

```

```

CC properties of the membrane associated cytoskeleton.
CC -!- SUBUNIT: Heterodimer with BIN1. Binds SH3GLB1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Peripheral membrane protein associated with
the cytoplasmic surface of synaptic vesicles.
CC -!- TISSUE SPECIFICITY: Is abundant in the forebrain and cerebellum.
CC It is also found in the adrenal gland, anterior and posterior
pituitary.
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X60422; CAA42953.1; -.
CC FIR; S22700; S22700.
CC InterPro; IPR003005; Amphiphysin.
CC InterPro; IPR003017; Amphiphysin_1.
CC InterPro; IPR006632; BAR_dom.
CC InterPro; IPR004148; BAR_dom.
CC InterPro; IPR001452; SH3.
CC Pfam; PF03114; BAR; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR01251; AMPHIPHYSIN.
CC ProDom; PD003208; Amphiphysin_1; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00721; BAR; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
KW Cytoskeleton; SH3 domain; Coiled coil.
FT DOMAIN 10 84
FT COILED COIL (POTENTIAL).
FT DOMAIN 144 191
FT COILED COIL (POTENTIAL).
FT DOMAIN 609 682
FT SH3.
SQ SEQUENCE 682 AA; 75205 MW; 61617F494F38EB20 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 682;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13
Db 19 RAQEKVLQKLGKA 31

RESULT 5
AMPH_RAT
ID AMPH RAT STANDARD; PRT; 683 AA.
AC O08838;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amphiphysin.
GN AMPH OR AMPHL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Brain cortex;
RX MEDLINE=98009145; PubMed=9348539;
RA Wigge P., Kohler K., Vallis Y., Owen D., Hunt S.P., McMahon H.T.;
RT "Amphiphysin heterodimers: potential role in clathrin-mediated
endocytosis.";
RL Mol. Biol. Cell 8:2003-2015(1997).
CC -!- FUNCTION: May participate in mechanisms of regulated exocytosis in
synapses and certain endocrine cell types. May control the
properties of the membrane associated cytoskeleton (By

```



```

CC similarity).
CC -1- SUBUNIT: Heterodimer with BIN1. Binds SH3GLB1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Peripheral membrane protein associated with
CC the cytoplasmic surface of synaptic vesicles (By similarity).
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y13381; CAA73808.1; -.
CC InterPro: IPR003005; Amphiphysin.
CC InterPro: IPR003017; Amphiphysin_1.
CC InterPro: IPR006632; BAR.
CC InterPro: IPR004148; BAR_dom.
CC InterPro: IPR001452; SH3.
CC Pfam: PF03114; BAR; 1.
CC Pfam: PF00018; SH3; 1.
CC PRINTS: PR01251; AMPHIPHYSIN.
CC PRINTS: PR00452; SH3DOMAIN.
CC ProDom: PD003208; Amphiphysin_1; 1.
CC ProDom: PD000066; SH3; 1.
CC SMART: SM00721; BAR; 1.
CC SMART: SM00326; SH3; 1.
CC PROSITE: PS50002; SH3; 1.
CC Cytoskeleton; SH3 domain; Coiled coil.
CC PT DOMAIN 10 83 COILED COIL (POTENTIAL).
CC FT DOMAIN 144 191 COILED COIL (POTENTIAL).
CC FT DOMAIN 610 683 SH3.
CC SQ SEQUENCE 683 AA; 74877 MW; 7FBA4A9E5A1F6631 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 683;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RAQEKVLQKLGKA 13
Db 19 RAQEKVLQKLGKA 31

RESULT 6
ID AMPH_HUMAN STANDARD; PRT; 695 AA.
AC P49418; O43538;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amphiphysin.
GN AMPH OR AMPHI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX MEDLINE=94357284; PubMed=8076697;
RA David C., Solimena M., de Camilli P.;
RT "Autoimmunity in stiff-man syndrome with breast cancer is targeted to
RT the C-terminal region of human amphiphysin, a protein similar to the
RT yeast proteins Rvs167 and Rvs161."
RL PNAS Lett. 351:73-79(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=95276740; PubMed=7757077;
RA Yamamoto R., Li X., Winter S., Francke U., Killmann M.W.;
RT "Primary structure of human amphiphysin, the dominant autoantigen of

```

```

RT paraneoplastic stiff-man syndrome, and mapping of its gene (AMPH) to
RT chromosome 7p13-p14."
RL Hum. Mol. Genet. 4:265-268(1995).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Breast;
RX MEDLINE=98174372; PubMed=9513187;
RA Floyd S.R., Butler M.H., Cremona O., David C., Freyberg Z., Zhang X.,
RA Solimena M., Tokunaga A., Ishizu H., Tautsui K., De Camilli P.V.;
RT "Expression of amphiphysin I, an autoantigen of paraneoplastic
RT neurological syndromes, in breast cancer."
RL Mol. Med. 4:29-39(1998).
CC -1- FUNCTION: May participate in mechanisms of regulated exocytosis in
CC synapses and certain endocrine cell types. May control the
CC properties of the membrane associated cytoskeleton.
CC -1- SUBUNIT: Heterodimer of AMPH and AMPHL (By similarity).
CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE CYTOPLASMIC SURFACE OF
CC SYNAPTIC VESICLES.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=128 kDa;
CC IsoId=P49418-1; Sequence=Displayed;
CC Name=2; Synonyms=108 kDa;
CC IsoId=P49418-2; Sequence=VSP_000245;
CC -1- TISSUE SPECIFICITY: Neurons, certain endocrine cell types and
CC spermatocytes.
CC -1- DISEASE: Patients with stiff-man syndrome, a rare disease of the
CC central nervous system characterized by progressive rigidity of
CC the body musculature with superimposed painful spasms, have
CC antibodies against AMPH.
CC -1- SIMILARITY: Contains 1 BAR domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U07616; AAA21865.1; -.
CC EMBL: X81438; CAA57197.1; -.
CC EMBL: AF034996; AAC02977.1; -.
CC PIR: S62400; S62400.
CC PDB: 1KY7; 12-JUN-02.
CC Genes: HGNC:471; AMPH.
CC MIM: 600418; -.
CC GO: GO:0015629; C:actin cytoskeleton; TAS.
CC GO: GO:0008021; C:synaptic vesicle; TAS.
CC GO: GO:0006897; P:endocytosis; TAS.
CC GO: GO:0007268; P:synaptic transmission; TAS.
CC InterPro: IPR003005; Amphiphysin.
CC InterPro: IPR003017; Amphiphysin_1.
CC InterPro: IPR006632; BAR.
CC InterPro: IPR004148; BAR_dom.
CC InterPro: IPR001452; SH3.
CC Pfam: PF03114; BAR; 1.
CC Pfam: PF00018; SH3; 1.
CC PRINTS: PR01251; AMPHIPHYSIN.
CC PRINTS: PR00452; SH3DOMAIN.
CC ProDom: PD003208; Amphiphysin_1; 1.
CC ProDom: PD000066; SH3; 1.
CC SMART: SM00721; BAR; 1.
CC SMART: SM00326; SH3; 1.
CC PROSITE: PS50002; SH3; 1.
CC Cytoskeleton; SH3 domain; Coiled coil; Alternative splicing;
CC 3D-structure.
CC DOMAIN 10 83 COILED COIL (POTENTIAL).
CC FT DOMAIN 144 191 COILED COIL (POTENTIAL);
CC FT DOMAIN 622 695 SH3.
CC VARSPLIC 425 466 Missing (in isoform 2).
CC /FtId=VSP_000245.

```

```

SQ SEQUENCE 695 AA; 76256 MW; 78B4F75AB75BA357 CRC64;
Query Match 100.0%; Score 61; DB 1; Length 695;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 13
Db 19 RAQEKVLQKLGK 31

RESULT 7
Y310_PSEAE
AC Q51470; STANDARD; PRT; 446 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein PA3980.
GN PA3980.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=96004705; PubMed=7565600;
RA Hungerer C., Troup B., Romling U., Jahn D.;
RT "Cloning, mapping and characterization of the Pseudomonas aeruginosa
hemL gene.";
RL Mol. Gen. Genet. 248:375-380(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- SIMILARITY: Belongs to the UPF0004 family.
CC -!- SIMILARITY: Contains 1 TRAM domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X82072; CAA57577.1;
DR EMBL; AE004816; AAG07367.1;
DR PIR; F83147; F83147.
DR PIR; S57900; S57900.
DR InterPro; IPR006638; E1p3.
DR InterPro; IPR006463; MiaB_methylase.
DR InterPro; IPR007197; Radical_SAM.
DR InterPro; IPR002792; TRAM.
DR InterPro; IPR005839; UPF0004.
DR Pfam; PF04055; Radical_SAM; 1.
DR Pfam; PF01938; TRAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR SMART; SM00729; E1p3; 1.
DR TIGRFAMs; TIGR01574; miaB_methylase; 1.
DR TIGRFAMs; TIGR00089; TIGR00089; 1.
DR PROSITE; PS50926; TRAM; 1.
DR PROSITE; PS01278; UPF0004; 1.

KW Hypothetical protein; Complete proteome.
FT DOMAIN 378 442 TRAM.
FT CONFLICT 55 60 A -> V (IN REF. 1).
FT CONFLICT 60 60 F -> C (IN REF. 1).
FT CONFLICT 395 395 K -> E (IN REF. 1).
SQ SEQUENCE 446 AA; 49978 MW; 9F5EA90E6288850 CRC64;

Query Match 65.6%; Score 40; DB 1; Length 446;
Best Local Similarity 72.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLG 11
Db 54 KAQEKVFSKLG 64

RESULT 8
CP72_CATRO
ID CP72_CATRO STANDARD; PRT; 524 AA.
AC Q05047;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 72A1 (EC 1.3.3.9) (CYP72A1) (Secologanin synthase)
DE (SLS).
GN CYP72A1 OR CYP72 OR P450CR3.
OS Catharanthus roseus (Rosa periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae;
OC Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. CP3A;
RA Vetter H.-P., Mangold U., Schroeder G., Marner F.-J.,
RA Werck-Reichhart D., Schroeder J.;
RT "Molecular analysis and heterologous expression of an inducible
cytochrome P-450 protein from periwinkle (Catharanthus roseus L.).";
RL Plant Physiol. 100:998-1007(1992).
RN [2]
RP SEQUENCE OF 469-524 FROM N.A.
RC STRAIN=cv. G. Don;
RX MEDLINE=93283641; PubMed=8507838;
RA Meijer A.H., Souer E., Verpoorte R., Hoge J.H.C.;
RT "Isolation of cytochrome P-450 cDNA clones from the higher plant
Catharanthus roseus by a PCR strategy.";
RL Plant Mol. Biol. 22:379-383(1993).
RN [3]
RP FUNCTION.
RC STRAIN=cv. CP3A;
RX MEDLINE=20575722; PubMed=11135113;
RA Imler S., Schroeder G., St-Pierre B., Crouch N.P., Hotze M.,
RA Schmidt J., Strack D., Matern U., Schroeder J.;
RT "Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme
activities and identification of cytochrome P450 CYP72A1 as
secologanin synthase";
RL Plant J. 24:797-804(2000).
CC -!- FUNCTION: Converts loganin into secologanin.
CC -!- CATALYTIC ACTIVITY: Loganin + NADPH + O(2) = secologanin + NADP(+)
+ 2 H(2)O.
CC -!- PATHWAY: Indole alkaloids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Vacuolar membrane (Probable).
CC -!- TISSUE SPECIFICITY: Upper and lower leaf epidermis.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
```

```

CC EMBL; LI0081; AAA33106.1; -
DR EMBL; XG9775; CAA49430.1; -
DR PIR; S35168; S35168.
DR PIR; T09944; T09944.
DR HSP; P14779; IUPZ.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;
KW NADP; Heme; Alkaloid metabolism.
FT TRANSMEM 12 32 POTENTIAL.
FT METAL 470 470 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT VARIANT 190 190 I -> L.
FT VARIANT 194 194 Q -> E.
FT VARIANT 223 223 E -> D.
FT VARIANT 312 312 K -> R.
FT VARIANT 318 318 S -> T.
FT VARIANT 403 403 V -> I.
FT VARIANT 405 405 K -> E.
FT VARIANT 411 411 S -> P.
SQ SEQUENCE 524 AA; 60557 MW; EP5D864E43C751E8 CRC64;
Query Match 63.9%; Score 39; DB 1; Length 524;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 RAQEKVLQKLGK 12
DB 357 RAREEVLQAFGK 368
||:|||||
||:|||||

RESULT 9
ARR2_ONCMY STANDARD; PRT; 405 AA.
AC P51467;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Arrestin, red cell isoform 2 (CARR).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96257743; PubMed=8687393;
RA Jahns R., Borgese F., Lindenthal S., Straub A., Motaïs R., Fievet B.;
RT "Trout red blood cell arrestin (TRCarr), a novel member of the
RT arrestin family: cloning, immunoprecipitation and expression of
RT recombinant TRCarr.";
RL Biochem. J. 316:497-506(1996).
CC -!- SIMILARITY: Belongs to the arrestin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U48410; AAB16954.1; -
DR PIR; S68253; S68253.
DR HSP; P08168; 1CF1.
DR InterPro; IPR000698; Arrestin.
DR Pfam; PF00339; arrestin; 1.
DR Pfam; PF02752; arrestin C; 1.
DR PRINTS; PR00309; ARRESTIN.
DR ProDom; PD002099; Arrestin; 1.
DR PROSITE; PS00295; ARRESTINS; 1.
KW Sensory transduction.
SQ SEQUENCE 407 AA; 45902 MW; 4986C2237E709F16 CRC64;
Query Match 62.3%; Score 38; DB 1; Length 407;
Best Local Similarity 59.3%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 RAQEKVLQKLGK 12
DB 100 RLQERLLKLGK 111
||:|||||
||:|||||

RESULT 11
ARR3_ONCMY STANDARD; PRT; 415 AA.
AC P51468;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

```

DR PROSITE; PS00295; ARRESTINS; 1.
KW Sensory transduction.
SQ SEQUENCE 405 AA; 45668 MW; B09B68171842F2E2 CRC64;
Query Match 62.3%; Score 38; DB 1; Length 405;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 RAQEKVLQKLGK 12
DB 100 RLQERLLKLGK 111
||:|||||
||:|||||

RESULT 10
ARR1_ONCMY STANDARD; PRT; 407 AA.
AC P51466;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Arrestin, red cell isoform 1 (CARR).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96257743; PubMed=8687393;
RA Jahns R., Borgese F., Lindenthal S., Straub A., Motaïs R., Fievet B.;
RT "Trout red blood cell arrestin (TRCarr), a novel member of the
RT arrestin family: cloning, immunoprecipitation and expression of
RT recombinant TRCarr.";
RL Biochem. J. 316:497-506(1996).
CC -!- SIMILARITY: Belongs to the arrestin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U48410; AAB16954.1; -
DR PIR; S68253; S68253.
DR HSP; P08168; 1CF1.
DR InterPro; IPR000698; Arrestin.
DR Pfam; PF00339; arrestin; 1.
DR Pfam; PF02752; arrestin C; 1.
DR PRINTS; PR00309; ARRESTIN.
DR ProDom; PD002099; Arrestin; 1.
DR PROSITE; PS00295; ARRESTINS; 1.
KW Sensory transduction.
SQ SEQUENCE 407 AA; 45902 MW; 4986C2237E709F16 CRC64;
Query Match 62.3%; Score 38; DB 1; Length 407;
Best Local Similarity 59.3%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 RAQEKVLQKLGK 12
DB 100 RLQERLLKLGK 111
||:|||||
||:|||||

RESULT 11
ARR3_ONCMY STANDARD; PRT; 415 AA.
AC P51468;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

DE Arrestin, red cell isoform 3 (CARR).

OS Oncorhynchus mykiss (Rainbow trout).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Telostei; Euteleostei;

OC Protactinopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

OX NCB1\_TaxID=8022;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=96257743; PubMed=8687393;

RX Jahns R., Borgese F., Lindenthal S., Straub A., Motais R., Fievet B.,

RT "Trout red blood cell arrestin (TRCarr), a novel member of the

RT arrestin family: cloning, immunoprecipitation and expression of

RT recombinant rCarr."

RL Biochem. J. 316:497-506 (1996).

CC -!- SIMILARITY: Belongs to the arrestin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U48412; AAB16956.1; -.

DR PIR; S68255; S68255.

DR HSP; P08168; ICF1.

DR InterPro; IPR000698; Arrestin.

DR InterPro; IPR007110; IG-like.

DR Pfam; PF00339; arrestin; 1.

DR Pfam; PF02752; arrestin; C. 1.

DR PRINTS; PR00309; ARRESTIN.

DR PRODOM; P000209; Arrestin; 1.

DR PROSITE; PS00295; ARRESTINS; 1.

DR Sensory transduction.

KW

SQ SEQUENCE 415 AA; 46770 MW; 7797256F1A2FBC8 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 415;

Best Local Similarity 58.3%; Pred. No. 36;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 RAQEKVLQKLGK 12

Db 100 RLQERLLKLGQ 111

RESULT 12

ARR2 BOVIN

ID ARR2 BOVIN STANDARD; PRT; 420 AA.

AC P32120; O77565; O77566;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Beta-arrestin 2 (Arrestin, beta 2) (Arrestin 3).

GN ARRB2.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCB1\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=93340166; PubMed=8340388;

RX Sterne-Marx R., Gurevich V.V., Goldsmith P., Bodine R.C., Sanders C.,

RA Doroso L.A., Benovic J.L.;

RT "Polypeptide variants of beta-arrestin and arrestin3."

RL J. Biol. Chem. 268:15640-15648 (1993).

RN [2]

RP SEQUENCE OF 304-396 FROM N.A.

RP TISSUE=Retina;

RC MEDLINE=99145674; PubMed=9767391;

RX Komori N., Cain S.D., Roch J.-M., Miller K.E., Matsumoto H.;

RT "Differential expression of alternative splice variants of beta-

RT arrestin-1 and -2 in rat central nervous system and peripheral

RT tissues."

RL Eur. J. Neurosci. 10:2607-2616 (1998).

CC -!- FUNCTION: Regulates beta-adrenergic receptor function. Beta-

CC arrestins seem to bind phosphorylated beta-adrenergic receptors,

CC thereby causing a significant impairment of their capacity to

CC activate G(s) proteins.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=2;

CC Name=Long;

CC IsoId=P32120-1; Sequence=Displayed;

CC Name=Short;

CC IsoId=P32120-2; Sequence=VSP\_000323;

CC -!- TISSUE SPECIFICITY: Found in a variety of tissues. The short

CC isoform is the most abundant form in all tissues.

CC -!- SIMILARITY: Belongs to the arrestin family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L14641; -; NOT ANNOTATED CDS.

DR EMBL; AF051456; AAC28615.1; -.

DR EMBL; AF051456; AAC28616.1; -.

DR PIR; A47140; A47140.

DR HSP; P08168; ICF1.

DR InterPro; IPR000698; Arrestin.

DR InterPro; IPR007110; IG-like.

DR Pfam; PF00339; arrestin; 1.

DR Pfam; PF02752; arrestin; C. 1.

DR PRINTS; PR00309; ARRESTIN.

DR PRODOM; P000209; Arrestin; 1.

DR PROSITE; PS00295; ARRESTINS; 1.

KW Sensory transduction; Nuclear protein; Alternative splicing.

FT DOMAIN 400 410 NUCLEAR EXPORT SIGNAL (BY SIMILARITY).

FT VARSPLIT 363 373 Missing (in isoform Short).

FT /FTID=VSP\_000323.

FT CONFLICT 362 362 A -> P (IN REF. 2).

FT SEQUENCE 420 AA; 47224 MW; 590ECF2D2D29F4D1 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 420;

Best Local Similarity 58.3%; Pred. No. 37;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 RAQEKVLQKLGK 12

Db 100 RLQERLLKLGQ 111

RESULT 13

DSL1 CHICK

ID DSL1 CHICK STANDARD; PRT; 427 AA.

AC P34822;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Dorsalin-1 precursor (DSL-1).

GN DSL-1.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCB1\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.; AND SEQUENCE OF 319-322.

RC TISSUE=Spinal cord;

RX MEDLINE=93272310; PubMed=7916656;

RA Basler K., Edlund T., Jessell T.M., Yamada T.;  
 RT "Control of cell pattern in the neural tube: regulation of cell  
 RT differentiation by dorsalin-1, a novel TGF beta family member.";  
 RL Cell 73:689-702(1993).  
 CC -!- FUNCTION: APPEARS TO REGULATE CELL DIFFERENTIATION WITHIN THE  
 CC NEURAL TUBE. MAY REGULATE THE DIFFERENTIATION OF CELL TYPES ALONG  
 CC THE DORSOVENTRAL AXIS OF THE NEURAL TUBE, ACTING IN CONJUNCTION  
 CC WITH DISTINCT VENTRALIZING SIGNALS FROM THE NOTOCHORD AND FLOOR  
 CC PLATE. CONTROLS THE CELL DIFFERENTIATION IN THE NEURAL TUBE IN  
 CC SEVERAL WAYS: (1) PROMOTES THE DIFFERENTIATION OF CELL TYPES THAT  
 CC DERIVE FROM THE DORSAL NEURAL TUBE. (2) ENSURES THAT THE DORSAL  
 CC NEURAL TUBE IS REFRACTORY TO VENTRALIZING SPECIES FROM THE  
 CC NOTOCHORD. (3) CAN DIFFUSE AND INFLUENCE THE FATE OF CELLS IN MORE  
 CC VENTRAL REGIONS OF THE NEURAL TUBE.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED SELECTIVELY IN THE DORSAL NEURAL  
 CC TUBE. LOWER LEVELS SEEN IN KIDNEY AND MYOTOMAL CELLS.  
 CC -!- DEVELOPMENTAL STAGE: IS NOT EXPRESSED IN NEURAL CELLS AT STAGES  
 CC BEFORE NEURAL TUBE CLOSURE. IS EXPRESSED AT HIGH LEVELS IN THE  
 CC DORSAL THIRD OF THE NEURAL TUBE, BEGINNING AT THE TIME OF NEURAL  
 CC TUBE CLOSURE, BUT NOT BY VENTRAL NEURAL CELLS OR BY NONNEURAL  
 CC CELLS. DORSAL RESTRICTION PERSISTS IN THE SPINAL CORD AT STAGES  
 CC AFTER THE ONSET OF NEURONAL DIFFERENTIATION. AT LATER STAGES OF  
 CC SPINAL DEVELOPMENT, IS RESTRICTED TO THE DORSOMEDIAL REGION OF  
 CC THE SPINAL CORD, INCLUDING BUT NOT CONFINED TO THE ROOF PLATE.  
 CC -!- SIMILARITY: Belongs to the TGF-beta family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L12032; AAA48752.1; -  
 CC F1R; A40735; A40735.  
 CC HSSP; P12643; 3BMP.  
 CC InterPro; IPR002405; Inhibin\_alpha.  
 CC InterPro; IPR001839; TGFb.  
 CC InterPro; IPR001111; TGFb\_N.  
 CC Pfam; PF00019; TGF-beta; 1.  
 CC Pfam; PF00688; TGFb\_propeptide; 1.  
 CC PRINTS; PR00669; INHIBIN.  
 CC ProDom; PD000357; TGFb; 1.  
 CC SMART; SM00204; TGFb; 1.  
 CC PROSITE; PS00250; TGF BETA 1; 1.  
 CC Signal; Growth factor; Cytokine; Glycoprotein.  
 CC SIGNAL 1 20 POTENTIAL.  
 CC PROPEP 21 318  
 CC CHAIN 319 427 DORSALIN-1.  
 CC FT DISULFID 325 391 BY SIMILARITY.  
 CC FT DISULFID 354 424 BY SIMILARITY.  
 CC FT DISULFID 358 426 BY SIMILARITY.  
 CC FT DISULFID 390 390 INTERCHAIN (BY SIMILARITY).  
 CC FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 265 285 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 427 AA; 48626 MW; 23AA42DC7085FABC CRC64;  
 SQ  
 Query Match 62.3%; Score 38; DB 1; Length 427;  
 Best Local Similarity 80.0%; Pred. No. 37;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 FT PROPEP 21 318  
 FT CHAIN 319 427 DORSALIN-1.  
 FT FT DISULFID 325 391 BY SIMILARITY.  
 FT FT DISULFID 354 424 BY SIMILARITY.  
 FT FT DISULFID 358 426 BY SIMILARITY.  
 FT FT DISULFID 390 390 INTERCHAIN (BY SIMILARITY).  
 FT FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT FT CARBOHYD 265 285 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 427 AA; 48626 MW; 23AA42DC7085FABC CRC64;  
 Query Match 62.3%; Score 38; DB 1; Length 427;  
 Best Local Similarity 80.0%; Pred. No. 37;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 QEKVLQKLGK 12  
 DB 282 QESVLNKLK 291  
 RESULT 14  
 Y426\_BUCAP

ID Y426\_BUCAP STANDARD; PRT; 439 AA.  
 AC Q8X9C2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 GN BUG426.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=98794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22084549; PubMed=12089438;  
 RA Tanas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,  
 RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;  
 RT "50 million years of genomic stasis in endosymbiotic bacteria.";  
 RL Science 296:2376-2379(2002).  
 CC -!- SIMILARITY: BELONGS TO THE UPF0004 FAMILY. STRONG, TO E.COLI YLEA.  
 CC -!- SIMILARITY: Contains 1 TRAM domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE014118; AM67969.1; -  
 CC InterPro; IPR006638; Elp3.  
 CC InterPro; IPR006463; MiaB\_methylase.  
 CC InterPro; IPR007197; Radical\_SAM.  
 CC InterPro; IPR002792; TRAM.  
 CC InterPro; IPR005839; UPF0004.  
 CC Pfam; PF04055; Radical\_SAM; 1.  
 CC Pfam; PF01938; TRAM; 1.  
 CC Pfam; PF00919; UPF0004; 1.  
 CC SMART; SM00729; Elp3; 1.  
 CC TIGRFAMs; TIGR01574; miaB\_methylase; 1.  
 CC TIGRFAMs; TIGR00089; TIGR00089; 1.  
 CC PROSITE; PS0926; TRAM; 1.  
 CC PROSITE; PS01278; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 FT DOMAIN 377 439 TRAM.  
 SQ SEQUENCE 439 AA; 50787 MW; 2E83E7F58CD0E5B8 CRC64;  
 Query Match 62.3%; Score 38; DB 1; Length 439;  
 Best Local Similarity 58.3%; Pred. No. 38;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKLGK 12  
 DB 53 KAQEKVFLQGR 64  
 RESULT 15  
 Y441\_BUCAI  
 ID Y441\_BUCAI STANDARD; PRT; 439 AA.  
 AC P57516;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 GN BU441.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tokyo 1998;

```

RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- SIMILARITY: BELONGS TO THE UPF0004 FAMILY. STRONG, TO E.COLI YLEA.
CC -!- SIMILARITY: Contains 1 TRAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF001119; BAB13139.1; -.
CC InterPro; IPR006638; EIP3.
CC InterPro; IPR006463; MiaB_methylase.
CC InterPro; IPR007197; Radical_SAM.
CC InterPro; IPR002792; TRAM.
CC InterPro; IPR005839; UPF0004.
CC Pfam; PF04055; Radical_SAM; 1.
CC Pfam; PF01938; TRAM; 1.
CC Pfam; PF00919; UPF0004; 1.
CC SMART; SM00729; EIP3; 1.
CC TIGRFAMs; TIGR01574; miaB_methylase; 1.
CC TIGRFAMs; TIGR00089; TIGR00089; 1.
CC PROSITE; PS0926; TRAM; 1.
CC PROSITE; PS01278; UPF0004; 1.
CC Hypothetical protein; Complete proteome.
KW DOMAIN 377 439 TRAM.
FT DOMAIN 377 439
SQ SEQUENCE 439 AA; 50464 MW; 287361B42483D610 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 439;
Best Local Similarity 58.3%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12
Db 53 KAQEKVFHQLGR 64

RESULT 16
YA01_PASMU STANDARD; PRT; 474 AA.
AC Q9L699;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein PM1001.
GN PM1001.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RA Fuller T.E., Kennedy M.J., Lowery D.E.;
RT "Identification of Pasteurella multocida virulence genes in a
RT septicemic mouse model using signature-tagged mutagenesis.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L., Faustian M.L., Whittam T.S., Kapur V.;
RT "Complete genome sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: BELONGS TO THE UPF0004 FAMILY. STRONG, TO E.COLI YLEA.
CC -!- SIMILARITY: Contains 1 TRAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF237940; AAP68426.1; -.
CC EMBL; AE006140; AK03085.1; -.
CC InterPro; IPR006638; EIP3.
CC InterPro; IPR006463; MiaB_methylase.
CC InterPro; IPR007197; Radical_SAM.
CC InterPro; IPR002792; TRAM.
CC InterPro; IPR005839; UPF0004.
CC Pfam; PF04055; Radical_SAM; 1.
CC Pfam; PF01938; TRAM; 1.
CC Pfam; PF00919; UPF0004; 1.
CC SMART; SM00729; EIP3; 1.
CC TIGRFAMs; TIGR01574; miaB_methylase; 1.
CC TIGRFAMs; TIGR00089; TIGR00089; 1.
CC PROSITE; PS0926; TRAM; 1.
CC PROSITE; PS01278; UPF0004; 1.
CC Hypothetical protein; Complete proteome.
KW DOMAIN 378 441 TRAM.
FT DOMAIN 378 441
SQ SEQUENCE 474 AA; 53553 MW; 610A390EBF6D6629 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 474;
Best Local Similarity 58.3%; Pred. No. 41;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RAQEKVLQKLGK 12
Db 54 KAQEKVFHQLGR 65

RESULT 17
YLEA_ECOLI STANDARD; PRT; 474 AA.
ID YLEA_ECOLI
AC P77645;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ylea.
GN YLEA OR B0661 OR C0747 OR Z0810 OR ECS0699.
OS Escherichia coli.
OS Escherichia coli O6, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CPT073 / ATCC 700928;
RX MEDLINE=22398234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
```

```
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RN Nature 409:529-533 (2001).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H., Shinagawa K.,
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RN DNA Res. 8:11-22 (2001).
RN [6]
RN -!- SIMILARITY: BELONGS TO THE UPF0004 FAMILY. STRONG, TO H.INFLUENZA
RN H10019.
CC -!- SIMILARITY: Contains 1 TRAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000170; AAC73762.1; -
DR EMBL; U82598; AAB40863.1; -
DR EMBL; AE016757; AAN79220.1; -
DR EMBL; AE005245; AAG54994.1; -
DR EMBL; AP002552; BAB34122.1; -
DR PIR; C64801; C64801.
DR PIR; C90716; C90716.
DR EcoGene; EG13657; YleA.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR006463; MiaB_methylase.
DR InterPro; IPR007197; Rad1Gal_SAM.
DR InterPro; IPR002792; TRAM.
DR InterPro; IPR005839; UPF0004.
DR Pfam; PF04055; Radical_SAM; 1.
DR Pfam; PF01938; TRAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR01574; miaB-methylase; 1.
DR TIGRFAMs; TIGR00089; TIGR00089; 1.
DR PROSITE; PS50926; TRAM; 1.
DR PROSITE; PS01278; UPF0004; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 378 441
FT SEQUENCE 474 AA; 53663 MW; 49CBB93888F79D CRC64;
SQ
Query Match 62.3%; Score 38; DB 1; Length 474;
Best Local Similarity 58.3%; Pred. No. 41;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 RAQEKVLQKLGK 12
:|||||:
Db 54 KAQEKVFLQGR 65
```

```
RESULT 18
VLEA_HAEIN STANDARD; PRT; 474 AA.
ID YLEA_HAEIN
AC Q57163;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein H10019.
GN H10019
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Kerlavage A.R., Sulton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA McKenney K., Sutton R., Liu L.-I., Glodek A., Kelley J.M.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RN Science 269:496-512 (1995).
RL Science 269:496-512 (1995).
CC -!- SIMILARITY: BELONGS TO THE UPF0004 FAMILY. STRONG, TO E.COLI YLEA.
CC -!- SIMILARITY: Contains 1 TRAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32687; AAC21697.1; -
DR PIR; D64140; D64140.
DR TIGR; H10019; -
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR006463; MiaB_methylase.
DR InterPro; IPR007197; Radical_SAM.
DR InterPro; IPR002792; TRAM.
DR InterPro; IPR005839; UPF0004.
DR Pfam; PF04055; Radical_SAM; 1.
DR Pfam; PF01938; TRAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMs; TIGR01574; miaB-methylase; 1.
DR TIGRFAMs; TIGR00089; TIGR00089; 1.
DR PROSITE; PS50926; TRAM; 1.
DR PROSITE; PS01278; UPF0004; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 378 441
FT SEQUENCE 474 AA; 53640 MW; A6CEDFBAE16923D CRC64;
SQ
Query Match 62.3%; Score 38; DB 1; Length 474;
Best Local Similarity 58.3%; Pred. No. 41;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 RAQEKVLQKLGK 12
:|||||:
Db 54 KAQEKVFLQGR 65
```

```
RESULT 19
YF01_MYCTU
ID YF01_MYCTU STANDARD; PRT; 273 AA.
```

P71782;  
 01-NOV-1997 (Rel. 35, Created)  
 01-NOV-1997 (Rel. 35, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 Hypothetical protein Rv1501/MT1550/Wh1539.  
 Rv1501 OR MT1550 OR MT1539 OR Wh1539.  
 OS Mycobacterium tuberculosis, and  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773, 1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gass S., Barry C.E. III, Tekaija P.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayan L.A., Emdolaeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL J. Bacteriol. 184:5479-5490 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.bovis; STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,  
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Atkin R., Deggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
 RA "The complete genome sequence of Mycobacterium bovis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).  
 CC -1- SIMILARITY: TO M.TUBERCULOSIS RV3633.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z79701; CAB02015.1; -;  
 CC EMBL; AE007023; AAK45816.1; -;  
 CC EMBL; BX248339; CAD96206.1; -;  
 CC PIR; H70712; H70712.  
 CC TIGR; MT1550; -;  
 CC TubercuList; Rv1501; -;  
 CC InterPro; IPR008775; PhYH.  
 CC Pfam; PF05721; PhYH; 1.  
 CC Hypothetical protein; Complete proteome.  
 KW SEQUENCE 273 AA; 30906 MW; 1F4DEED028A2AE84 CRC64;  
 Query Match 60.7%; Score 37; DB 1; Length 273;  
 Best Local Similarity 50.0%; Pred. No. 37;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
 DB 45 RVQERILTEIGK 56  
 RESULT 20  
 ARGCG METJJA STANDARD; PRT; 341 AA.  
 AC Q58456;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-  
 DE acetyl-L-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).  
 GN ARGCG OR MT1096.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kervilave A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073 (1996).  
 CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+) +  
 CC phosphate = N-acetyl-5-glutamyl phosphate + NADPH.  
 CC -1- PATHWAY: Arginine biosynthesis; third step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: Belongs to the NAGSA dehydrogenase family. Subfamily  
 CC 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U67552; AAB99039.1; ALT\_INIT.  
 CC TIGR; MJ1036; -;  
 CC HAMAP; MF\_00150; -; 1.  
 CC InterPro; IPR000706; AGPR\_act\_site.  
 CC InterPro; IPR000534; Semialdh\_dh.  
 CC Pfam; PF01118; Semialdehyde\_dh; 1.  
 CC Pfam; PF02774; Semialdehyde\_dhC; 1.  
 CC ProDom; PD003765; AGPR\_act\_site; 1.  
 CC PROSITE; PS01224; ARGCG; 1.  
 CC Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.  
 KW Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.  
 FT ACT SITE 149 149 BY SIMILARITY.  
 SQ SEQUENCE 341 AA; 38172 MW; 72AE13AD35061D88 CRC64;  
 Query Match 60.7%; Score 37; DB 1; Length 341;  
 Best Local Similarity 80.0%; Pred. No. 45;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EKVLQKLGKA 13  
 DB 214 EKELKLGKA 223  
 RESULT 21  
 P35\_MYCPE



ID P35 MYCPE STANDARD; PRT; 362 AA.  
 AC Q50367;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Lipoprotein p35 precursor.  
 GN MPE810.  
 OS Mycoplasma penetrans.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OC NCBI TaxID=28227;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 211-225.  
 RC STRAIN=GTU-54-6A1;  
 RX MEDLINE=95377619; PubMed=7649455;  
 RA Ferris S., Watson H.L., Neyrolles O., Montagnier L., Blanchard A.;  
 RT "Characterization of a major Mycoplasma penestrans lipoprotein and of  
 its gene.";  
 RL FEMS Microbiol. Lett. 130:313-319(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HF-2;  
 RX MEDLINE=22354719; PubMed=12466555;  
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,  
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;  
 RT "The complete genomic sequence of Mycoplasma penestrans, an  
 intracellular bacterial pathogen in humans.";  
 RL Nucleic Acids Res. 30:5293-5300(2002).  
 RN [3]  
 RP CHARACTERIZATION.  
 RC STRAIN=GTU-54-6A1;  
 RX MEDLINE=99173739; PubMed=10075417;  
 RA Neyrolles O., Eliane J.-P., Ferris S., da Cunha R.A.F., Prevost M.-C.,  
 RA Bahraoui E., Blanchard A.;  
 RT "Antigenic characterization and cytolocalization of P35, the major  
 Mycoplasma penestrans antigen.";  
 RL Microbiology 145:343-355(1999).  
 CC -!- FUNCTION: Major P35-penetrans antigen.  
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor.  
 CC Distributed all over the plasma membrane.  
 CC -!- PTM: The N-terminus is blocked (Probable).  
 CC -!- MISCELLANEOUS: The dominant B-epitopes are found at the N- and C-  
 terminal regions.  
 CC -!- SIMILARITY: BELONGS TO THE LIPOPROTEIN P35 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; I38250; AAC16392.1; -;  
 DR EMBL; AP004172; BAC44473.1; -;  
 DR InterPro; IPR000437; Prok\_lipprot\_S.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Antigen; Lipoprotein; Membrane; Signal; Complete proteome; Palmitate.  
 FT SIGNAL 1 30 PROBABLE.  
 FT CHAIN 31 362 LIPOPROTEIN P35.  
 FT LIPID 31 31 N-palmitoyl cysteine (Probable).  
 FT LIPID 31 31 S-diacylglycerol cysteine (Probable).  
 FT CONFLICT 354 354 D -> E (IN REF. 1).  
 SQ SEQUENCE 362 AA; 38370 MW; 6260A719B76F8937 CRC64;  
 Query Match 60.7%; Score 37; DB 1; Length 362;  
 Best Local Similarity 80.0%; Pred. No. 48;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 EKVLQKLGKA 13  
 |||||  
 DB 281 EKVLKLGYA 290

RESULT 22  
 PCP1 SCHPO STANDARD; PRT; 1208 AA.  
 AC Q92351;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Spindle pole body protein pcpl1.  
 GN PCP1 OR SPAC6G9.06C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OC NCBI TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
 RC MEDLINE=21852775; PubMed=11864908;  
 RA Flory M.R., Morpew M., Joseph J.D., Means A.R., Davis T.N.;  
 RT "Pcp1p, a Spcl10p-related calmodulin target at the centrosome of the  
 fission yeast Schizosaccharomyces pombe.";  
 RL Cell Growth Differ. 13:47-58(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy I., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
 RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Petashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Spindle pole body component that binds calmodulin.  
 CC Overexpression of pcpl causes the formation of supernumerary SPB-  
 like structures and disrupts both mitotic spindle assembly and  
 chromosome segregation.  
 CC -!- SUBCELLULAR LOCATION: Spindle pole body.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; Z81317; CAB03608.1; -;  
 DR EMBL; AF348506; AAK31344.1; -;  
 DR PIR; T39068; T39068.  
 DR GeneDB Spombe; SPAC6G9.06c; -;  
 KW Calmodulin-binding; Coiled coil.  
 FT DOMAIN 151 375 COILED COIL (POTENTIAL).  
 FT DOMAIN 387 803 COILED COIL (POTENTIAL).  
 FT DOMAIN 874 1091 COILED COIL (POTENTIAL).

FT DOMAIN 1177 1204 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 1208 AA; 140763 MW; 70264159ADD42424 CRC64;

Query Match  
 Best Local Similarity 58.3%; Score 37; DB 1; Length 1208;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RAQKVLOKLGK 12  
 | | | | | | | | | |  
 DB 152 REQKVLEKVR 163

RESULT 23  
 CHD1\_YEAST  
 ID CHD1\_YEAST STANDARD; PRT; 1468 AA.  
 AC P32657;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE CHD1 protein.  
 GN CHD1 OR YER164W OR SYGP-ORF4.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=97313264; PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
 RA Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
 RL Nature 387:78-81(1997).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential)  
 CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.  
 CC -!- SIMILARITY: Contains 2 chromo domains.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; U18917; AAB64691.1; -;  
 PIR; S30818; S30818.  
 DR GenOnline; I39241; -;  
 DR SGD; S0000966; CHD1.  
 DR GO; GO:0008023; C:transcription elongation factor complex; IPI.  
 DR GO; GO:0016887; F:ATPase activity; IDA.  
 DR GO; GO:0016944; F:Pol II transcription elongation factor acti. . .; IPI.  
 DR GO; GO:0006338; P:chromatin remodeling; IDA.  
 DR GO; GO:0006368; P:RNA elongation from Pol II promoter; IPI.  
 DR InterPro; IPR000953; Chromo.  
 DR InterPro; IPR001410; DRAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR000330; SNF2\_N.  
 DR Pfam; PF00385; chromo; 2.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00249; myb DNA-binding; 1.  
 DR Pfam; PF00176; SNF2\_N; 1.  
 DR SMART; SM00258; CHROMO; 2.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00450; HELIC\_C; 1.  
 DR PROSITE; PS00598; CHROMO\_1; 2.  
 DR PROSITE; PS00013; CHROMO\_2; 2.  
 DR PROSITE; PS00690; DEAH\_ATP\_HELICASE; FALSE\_NEG.

KW Hydrolase; Helicase; Nuclear protein; ATP-binding; DNA-binding;  
 KW Repeat.  
 FT DOMAIN 195 257 CHROMO 1.  
 FT DOMAIN 285 350 CHROMO 2.  
 FT NP\_BIND 401 408 ATP (POTENTIAL).  
 FT SITE 513 516 DEAH BOX.  
 SQ SEQUENCE 1468 AA; 168240 MW; 788DB74C7FEC6BE5 CRC64;

Query Match  
 Best Local Similarity 60.7%; Score 37; DB 1; Length 1468;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 AQEKVLOKLG 11  
 | | | | | | | | | |  
 DB 668 ABERVLOKFG 677

RESULT 24  
 HIS2\_METTH  
 ID HIS2\_METTH STANDARD; PRT; 96 AA.  
 AC O27344;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31) (PRA-PH).  
 GN HIS2 OR MTH1283.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 CC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 J. F. Bacteriol. 179:7135-7155(1997).  
 CC -!- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-ATP + H(2)O = 1-(5-  
 phosphoribosyl)-AMP + diphosphate.  
 CC -!- PATHWAY: Histidine biosynthesis; second step.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the PRA-PH family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE000893; AAB85765.1; -;  
 PIR; G69037; G69037.  
 DR HAMAP; MF 01020; -; 1.  
 DR InterPro; IPR008179; PRA-PH.  
 DR InterPro; IPR008178; PRA-PH/CH.  
 DR Pfam; PF01503; PRA-PH; 1.  
 DR ProDom; PD002611; PRA-PH/CH; 1.  
 DR Histidine biosynthesis; Hydrolase; Complete proteome.  
 SQ SEQUENCE 96 AA; 11122 MW; D95A7F8E0672CE8B CRC64;

Query Match  
 Best Local Similarity 59.0%; Score 36; DB 1; Length 96;  
 Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 2 AQEKVLOKLGK 12

```

Db 37 ABDKLEKICE 47
RESULT 25
REV_GINBI
ID REV_GINBI STANDARD; PRT; 155 AA.
AC Q9GFL3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S7.
GN RPS7.
OS Ginkgo biloba (Ginkgo).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Ginkgoales; Ginkgoaceae; Ginkgo.
OX NCBI_TaxID=3311;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11080123;
RA Graham S.W., Olmstead R.G.;
RT "Utility of 17 chloroplast genes for inferring the phylogeny of the
RT basal angiosperms";
RL Am. J. Bot. 87:1712-1730(2000).
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the head
CC domain of the 30S subunit (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF123779; AAG26116.1; -.
CC HSP; F17291; IRSS.
CC HAMAP; MF_00480; -.
CC InterPro; IPR000235; Ribosomal_S7.
CC InterPro; IPR005717; Ribosomal_S7_b/o.
CC Pfam; PF00177; Ribosomal_S7; 1.
CC ProDom; PD000817; Ribosomal_S7; 1.
CC TIGRfam; TIGR01029; rpsG_bact; 1.
CC PROSITE; PS00052; RIBOSOMAL_S7; FALSE NEG.
KW Ribosomal protein; rRNA-binding; rRNA-binding; Chloroplast.
SQ SEQUENCE 155 AA; 17782 MW; 18155A6FC446F369 CRC64;
Query Match 59.0%; Score 35; DB 1; Length 155;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 RAQKVKVLOKLGK 12
Db 45 RAMKLIQKTKG 56
RESULT 26
STAR_MESAU
ID STAR_MESAU STANDARD; PRT; 284 AA.
AC P70114;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Steroidogenic acute regulatory protein, mitochondrial precursor
DE (STAR) (STAR1).
GN STAR.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Adrenal gland;
RA Fleury A., Cloutier M., Ducharme L., Lefebvre A., LeHoux J.,
RA Lehoux J.-G.;
RT "Adrenocorticotropin regulates the level of the steroidogenic acute
RT regulatory (STAR) protein mRNA in hamster adrenals.";
RL Endocr. Res. 22:515-520(1996).
CC -!- FUNCTION: Plays a key role in steroid hormone synthesis by
CC enhancing the metabolism of cholesterol into pregnenolone.
CC Mediates the transfer of cholesterol from the outer mitochondrial
CC membrane to the inner mitochondrial membrane where it is cleaved
CC to pregnenolone (By similarity).
CC -!- PATHWAY: Steroidogenesis; rate-limiting step.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: Contains 1 START domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U66490; AAB06763.1; -.
CC HSP; Q14849; 1EM2.
CC InterPro; IPR000799; STAR.
CC InterPro; IPR002913; START.
CC Pfam; PF01852; START; 1.
CC PRINTS; PR00978; STARPROTEIN.
CC SMART; SM00234; START; 1.
CC PROSITE; PS50848; START; 1.
KW Lipid-binding; lipid transport; Transport; Steroidogenesis;
KW Mitochondrion; Transit peptide; Phosphorylation.
FT TRANSIT 1
FT CHAIN ? 284
FT DOMAIN 66 279
FT MOD RES 56 56
FT MOD RES 194 194
FT SEQUENCE 284 AA; 31497 MW; F23151F9EC8512B CRC64;
Query Match 59.0%; Score 36; DB 1; Length 284;
Best Local Similarity 87.5%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 5 KVLQKLGK 12
Db 154 KVLQKLGK 161
RESULT 27
STAR_BOVIN
ID STAR_BOVIN STANDARD; PRT; 285 AA.
AC Q28918;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Steroidogenic acute regulatory protein, mitochondrial precursor
DE (STAR) (STAR1).
GN STAR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Corpus luteum;

```

RX MEDLINE=96011827; PubMed=7488004;  
RA Hartung S., Rust W., Balvers M., Ivell R.;  
RT "Molecular cloning and in vivo expression of the bovine steroidogenic  
RL acute regulatory protein.";  
RL Biochem. Biophys. Res. Commun. 215:646-653 (1995).  
RN [2]  
RP REVISIONS TO 85 AND 174.  
RA Hartung S., Rust W., Balvers M., Ivell R.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RX PHOSPHORYLATION.  
RA MEDLINE=95352521; PubMed=7626524;  
RA Hartigan J.A., Green E.G., Mortensen R.M., Menachery A.,  
RA Williams G.H., Orme-Johnson N.R.;  
RT "Comparison of protein phosphorylation patterns produced in adrenal  
RT cells by activation of cAMP-dependent protein kinase and Ca-dependent  
RT protein kinase.";  
RL J. Steroid Biochem. Mol. Biol. 53:95-101 (1995).  
RN [3]  
CC -!- FUNCTION: Plays a key role in steroid hormone synthesis by  
CC enhancing the metabolism of cholesterol into pregnenolone.  
CC Mediates the transfer of cholesterol from the outer mitochondrial  
CC membrane to the inner mitochondrial membrane where it is cleaved  
CC to pregnenolone (By similarity).  
CC -!- PATHWAY: Steroidogenesis; rate-limiting step.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
CC -!- TISSUE SPECIFICITY: Corpus luteum and adrenal gland.  
CC -!- SIMILARITY: Contains 1 START domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Y17259; CAA76717.1; -  
CC EMBL; Y17260; CAA76718.1; -  
CC PIR; J43315; JC4315.  
CC HSP; Q14849; 1EW2.  
CC InterPro; IPR000799; STAR.  
CC InterPro; IPR002913; START.  
CC Pfam; PF01852; START; 1.  
CC PRINTS; PR00978; STARPROTEIN.  
CC SMART; SM00234; START; 1.  
CC PROSITE; PS50848; START; 1.  
CC Lipid-binding; Lipid transport; Transport; Steroidogenesis;  
KW Mitochondrion; Transit peptide; Phosphorylation.  
FT TRANSIT 1 ? 285 MITOCHONDRIAL ACUTE REGULATORY PROTEIN.  
FT CHAIN ? 285 START.  
FT DOMAIN 67 280  
FT MOD\_RES 57 57 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT MOD\_RES 195 195 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT CONFLICT 75 75 H -> Y (IN REF. 1; CAA76718).  
FT CONFLICT 99 99 E -> K (IN REF. 1; CAA76718).  
SQ SEQUENCE 285 AA; 31845 MW; EAFD5C14EE95E235 CRC64;  
Query Match 59.0%; Score 36; DB 1; Length 285;  
Best Local Similarity 87.5%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KVLQKLGK 12  
DB 155 KVLQKLGK 162  
RESULT 28  
STAR\_HORSE  
ID STAR\_HORSE STANDARD; PRT; 285 AA.  
AC O46689;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Steroidogenic acute regulatory protein, mitochondrial precursor  
DE (STAR) (StARD1).  
GN STAR.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pollicle;  
RX MEDLINE=99124482; PubMed=9927292;  
RA Kerban A., Boerboom D., Sirois J.;  
RT "Human chorionic gonadotropin induces an inverse regulation of  
RT steroidogenic acute regulatory protein messenger ribonucleic acid in  
RT theca interna and granulosa cells of equine preovulatory follicles.";  
RL Endocrinology 140:667-674 (1999).  
RN [1]  
CC -!- FUNCTION: Plays a key role in steroid hormone synthesis by  
CC enhancing the metabolism of cholesterol into pregnenolone.  
CC Mediates the transfer of cholesterol from the outer mitochondrial  
CC membrane to the inner mitochondrial membrane where it is cleaved  
CC to pregnenolone (By similarity).  
CC -!- PATHWAY: Steroidogenesis; rate-limiting step.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
CC -!- SIMILARITY: Contains 1 START domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF031697; AAC04704.1; -  
CC EMBL; AF031696; AAC04703.1; -  
CC HSP; Q14849; 1EW2.  
CC InterPro; IPR000799; STAR.  
CC InterPro; IPR002913; START.  
CC Pfam; PF01852; START; 1.  
CC PRINTS; PR00978; STARPROTEIN.  
CC SMART; SM00234; START; 1.  
CC PROSITE; PS50848; START; 1.  
CC Lipid-binding; Lipid transport; Transport; Steroidogenesis;  
KW Mitochondrion; Transit peptide; Phosphorylation.  
FT TRANSIT 1 ? 285 MITOCHONDRIAL ACUTE REGULATORY PROTEIN.  
FT CHAIN ? 285 START.  
FT DOMAIN 67 280  
FT MOD\_RES 57 57 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
FT MOD\_RES 195 195 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
SQ SEQUENCE 285 AA; 31852 MW; B0360EF93399727D CRC64;  
Query Match 59.0%; Score 36; DB 1; Length 285;  
Best Local Similarity 87.5%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KVLQKLGK 12  
DB 155 KVLQKLGK 162  
RESULT 29  
STAR\_HUMAN  
ID STAR\_HUMAN STANDARD; PRT; 285 AA.  
AC P49675; Q16396;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Steroidogenic acute regulatory protein, mitochondrial precursor  
DE (STAR) (StARD1).  
GN STAR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



DR HSSP; Q14849; 1EM2.  
 DR Genew; HGNC:11359; STAR.  
 DR MIN; 600617; --  
 DR MIN; 201710; --  
 DR GO; GO:0006694; P:steroid biosynthesis; TAS.  
 DR InterPro; IPR000799; STAR.  
 DR InterPro; IPR002913; START.  
 DR Pfam; PF01852; START; 1.  
 DR PRINTS; PR00978; STARPROTEIN.  
 DR SMART; SM00234; START; 1.  
 DR PROSITE; PS00848; START; 1.  
 KW Lipid-binding; Lipid transport; Transport; Steroidogenesis;  
 KW Mitochondrion; Transit peptide; Phosphorylation; Disease mutation;  
 KW Polymorphism.  
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).  
 FT CHAIN ? 285 STEROIDOGENIC ACUTE REGULATORY PROTEIN.  
 FT DOMAIN 67 START.  
 FT MOD\_RES 57 PHOSPHORYLATION (BY PKA).  
 FT MOD\_RES 195 PHOSPHORYLATION (BY PKA).  
 FT VARIANT 169 E -> G (in CLAH; partial loss of activity).  
 FT FTID=VAR\_014236.  
 FT VARIANT 169 E -> K (in CLAH; partial loss of activity).  
 FT FTID=VAR\_014237.  
 FT VARIANT 182 R -> L (in CLAH; partial loss of activity).  
 FT FTID=VAR\_005627.  
 FT VARIANT 203 D -> A.  
 FT FTID=VAR\_005628.  
 FT VARIANT 217 R -> T (in CLAH).  
 FT FTID=VAR\_014238.  
 FT VARIANT 218 A -> V (in CLAH; partial loss of activity).  
 FT FTID=VAR\_014239.  
 FT VARIANT 225 M -> T (in CLAH).  
 FT FTID=VAR\_014240.  
 FT VARIANT 272 Missing (in CLAH; partial loss of activity).  
 FT FTID=VAR\_014241.  
 FT VARIANT 275 L -> P (in CLAH; partial loss of activity).  
 FT FTID=VAR\_014242.  
 Query Match 59.0%; Score 36; DB 1; Length 285;  
 Best Local Similarity 87.5%; Pred. No. 58;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 KVLQKLGK 12  
 Db 155 KVLQKLGK 162  
 RESULT 30  
 STAR SHEEP  
 ID STAR SHEEP STANDARD; PRT; 285 AA.  
 AC P79245; O9GMD0;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Steroidogenic acute regulatory protein, mitochondrial precursor  
 DE (STAR) (STAR1).  
 OS STAR.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OC NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21066334; PubMed=11145616;  
 RA West L.A., Horvat R.D., Rees D.A., Barisas B.G., Juengel J.L.,  
 Niswender G.D.;

"Steroidogenic acute regulatory protein and peripheral-type benzodiazepine receptor associate at the mitochondrial membrane.";  
 Endocrinology 142:502-505(2001).  
 [2]  
 RP SEQUENCE OF 97-217 FROM N.A.  
 RC TISSUE=Corpus luteum;  
 RX MEDLINE=96079887; PubMed=7588291;  
 RA Juengel J.L., Meberg B.M., Turzillo A.M., Nett T.M., Niswender G.D.;  
 RT "Hormonal regulation of messenger ribonucleic acid encoding steroidogenic acute regulatory protein in ovine corpora lutea.";  
 RT Endocrinology 136:5423-5429(1995).  
 CC -1- FUNCTION: Plays a key role in steroid hormone synthesis by enhancing the metabolism of cholesterol into pregnenolone. Mediates the transfer of cholesterol from the outer mitochondrial membrane to the inner mitochondrial membrane where it is cleaved to pregnenolone (By similarity).  
 CC -1- PATHWAY: Steroidogenesis; rate-limiting step.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: Contains 1 START domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF290202; AAG02464.1; --  
 DR EMBL; S80098; AAB47088.1; --  
 DR HSSP; Q14849; 1EM2.  
 DR InterPro; IPR000799; STAR.  
 DR InterPro; IPR002913; START.  
 DR Pfam; PF01852; START; 1.  
 DR PRINTS; PR00978; STARPROTEIN.  
 DR SMART; SM00234; START; 1.  
 DR PROSITE; PS00848; START; 1.  
 KW Lipid-binding; Lipid transport; Transport; Steroidogenesis;  
 KW Mitochondrion; Transit peptide; Phosphorylation.  
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).  
 FT CHAIN ? 285 STEROIDOGENIC ACUTE REGULATORY PROTEIN.  
 FT DOMAIN 67 280 START.  
 FT MOD\_RES 57 57 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 FT MOD\_RES 195 195 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 FT CONFLICT 100 100 S -> N (IN REF. 2).  
 FT CONFLICT 103 103 V -> A (IN REF. 2).  
 FT CONFLICT 186 186 S -> R (IN REF. 2).  
 SQ SEQUENCE 285 AA; 31895 MW; AB130FD90A1DF295 CRC64;  
 Query Match 59.0%; Score 36; DB 1; Length 285;  
 Best Local Similarity 87.5%; Pred. No. 58;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 KVLQKLGK 12  
 Db 155 KVLQKLGK 162  
 RESULT 31  
 ERA THETN  
 ID ERA THETN STANDARD; PRT; 298 AA.  
 AC Q9RB50;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE GRP-binding protein era homolog.  
 DE ERA OR TTE0974.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OC NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity
CC and is essential for cell growth (By similarity).
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC proteins. Era subfamily.
CC -!- SIMILARITY: Contains 1 KH type-2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB013062; AAM24229.1; -
CC HAMAP; MF_00367; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR005662; Era.
CC InterPro; IPR005289; GTP-binding_dom.
CC InterPro; IPR006073; GTP1_OBG.
CC InterPro; IPR009019; KH_Prot.
CC InterPro; IPR004044; KH_TYPE_2.
CC InterPro; IPR002917; MHR_HSR1.
CC InterPro; IPR005225; Small GTP.
CC Pfam; PF01926; MHR_HSR1; 1.
CC PRINTS; PR00326; GTP1_OBG.
CC SMART; SM00382; AAA; 1.
CC TIGRfams; TIGR00436; era; 1.
CC TIGRfams; TIGR00650; MG442; 1.
CC TIGRfams; TIGR00231; small GTP; 1.
CC PROSITE; PS00823; KH_TYPE_2; 1.
CC PROSITE; PS00823; KH_TYPE_2; 1.
CC GTP-binding; RNA-binding; Complete proteome.
CC NP_BIND 12 19 GTP (POTENTIAL).
CC NP_BIND 59 63 GTP (POTENTIAL).
CC NP_BIND 121 124 GTP (POTENTIAL).
CC DOMAIN 202 280 KH_TYPE-2.
CC SEQUENCE 298 AA; 34113 MW; 902699D9202AA8A1 CRC64;
CC -----
CC Query Match 59.0%; Score 36; DB 1; Length 298;
CC Best Local Similarity 61.5%; Pred. No. 60;
CC Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC Qy 1 RAQEKVLQKLGK 13
CC | | | | |
CC Db 283 RDNEKLRLKLGYA 295
CC -----
RRHH_HELVI
ID ARRH_HELVI STANDARD; PRT; 381 AA.
AC P55274;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Arrestin homolog.
OS Heliothis virescens (Noctuid moth) (Owllet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
CC NCBI_TaxID=7102;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE=Antenna;
CC MEDLINE=93199955; PubMed=8452755;
CC Raming K., Freitag J., Krieger J., Breer H.;
CC "Arrestin-subtypes in insect antennae.";

```

```

RL Cell. Signal. 5:69-80(1993).
CC -!- SIMILARITY: Belongs to the arrestin family.
CC PIR; B56607; B56607.
DR HSP; P08168; ICF1.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00339; arrestin; 1.
DR Pfam; PF02752; arrestin_C; 1.
DR PRINTS; PR00309; ARRESTIN.
DR ProDom; PD002099; Arrestin; 1.
DR PROSITE; PS00295; ARRESTINS; 1.
KW Sensory transduction.
SQ SEQUENCE 381 AA; 42747 MW; 84BB92B1BB3DA573 CRC64;
CC -----
CC Query Match 59.0%; Score 36; DB 1; Length 381;
CC Best Local Similarity 54.5%; Pred. No. 74;
CC Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC Qy 1 RAQEKVLQKLG 11
CC | | | | |
CC Db 98 RTQERLIKLG 108
CC -----
RESULT 33
ID ARRI_RABIT STANDARD; PRT; 410 AA.
AC Q95223;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-arrestin 1 (Arrestin, beta 1).
GN ARRB1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CC NCBI_TaxID=9986;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=New Zealand white;
CC Man X.L., Sears J.E., Sears M.L.;
CC Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Regulates beta-adrenergic receptor function. Beta-
CC arrestins seem to bind phosphorylated beta-adrenergic receptors,
CC thereby causing a significant impairment of their capacity to
CC activate G(s) proteins.
CC -----
CC -!- SIMILARITY: Belongs to the arrestin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U75838; AAC48753.1; -
CC HSP; P08168; ICF1.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00339; arrestin; 1.
DR Pfam; PF02752; arrestin_C; 1.
DR PRINTS; PR00309; ARRESTIN.
DR ProDom; PD002099; Arrestin; 1.
DR PROSITE; PS00295; ARRESTINS; 1.
KW Sensory transduction.
SQ SEQUENCE 410 AA; 46360 MW; A8AB781CF089B65A CRC64;
CC -----
CC Query Match 59.0%; Score 36; DB 1; Length 410;
CC Best Local Similarity 50.0%; Pred. No. 79;
CC Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC Qy 1 RAQEKVLQKLGK 12
CC | | | | |

```

Db 99 RLQERLIKGLGE 110

Best Local Similarity 50.0%; Pred. NO. 81;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVQLKGLK 12  
DB 99 RLQERLIKGLGE 110

RESULT 34

ARBL\_BOVIN STANDARD; PRT; 418 AA.

AC P17870;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Beta-arrestin 1 (Arrestin, beta 1).  
GN ARRB1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOPROFORM 1A AND 1B).  
RX MEDLINE=90296080; PubMed=2163110;  
RA Lohse M.J., Benovic J.L., Codina J., Caron M.G., Lefkowitz R.J.;  
RT "Beta-arrestin: a protein that regulates beta-adrenergic receptor  
function.";  
RL Science 248:1547-1550 (1990).  
RN [2]  
RP ALTERNATIVE SPLICING.  
RC TISSUE=Brain;  
RX MEDLINE=93340166; PubMed=8340388;  
RA Sterne-Marr R., Gurevich V.V., Goldsmith P., Bodine R.C., Sanders C.,  
RA Donoso L.A., Benovic J.L.;  
RT "Polypeptide variants of beta-arrestin and arrestin3.";  
RL J. Biol. Chem. 268:15640-15648 (1993).  
CC -1- FUNCTION: Regulates beta-adrenergic receptor function. Beta-  
arrestins seem to bind phosphorylated beta-adrenergic receptors,  
thereby causing a significant impairment of their capacity to  
activate G(s) proteins.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1A;  
CC IsoId=P17870-1; Sequences=Displayed;  
CC Name=1B;  
CC IsoId=P17870-2; Sequences=VSP\_000321;  
CC -1- TISSUE SPECIFICITY: Beta-arrestin 1A is found in cortex,  
cerebellum, striatum, pineal gland, retina and heart. Beta-  
arrestin 1B is found in spleen, lung, pituitary and kidney.  
CC -1- SIMILARITY: Belongs to the arrestin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; M33601; AAA30387.1;  
DR EMBL; A34851; A34851.  
DR PDB; 1G4M; 03-OCT-01.  
DR PDB; 1G4R; 03-OCT-01.  
DR PDB; 1J5Y; 27-MAR-02.  
DR InterPro; IPR000698; Arrestin.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00339; arrestin; 1.  
DR Pfam; PF02752; arrestin; 1.  
DR PRINTS; PR00309; ARRESTIN.  
DR PRODOM; PD002099; ARRESTIN; 1.  
DR PROSITE; PS00295; ARRESTINS; 1.  
KW Sensory transduction; Alternative splicing; 3D-structure.  
FT VARSPLIC 334 341 Missing (in isoform 1B).  
FT /FTId=VSP\_000321.  
SQ SEQUENCE 418 AA; 47131 MW; 345302C20FA3360 CRC64;  
Query Match 59.0%; Score 36; DB 1; Length 418;

Best Local Similarity 50.0%; Pred. NO. 81;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQKVQLKGLK 12  
DB 99 RLQERLIKGLGE 110

RESULT 35

ARBL\_HUMAN STANDARD; PRT; 418 AA.

AC P49407; O75625; O75630; Q9BTK8;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Beta-arrestin 1 (Arrestin, beta 1).  
GN ARRB1 OR ARRB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOPROFORM 1A AND 1B).  
RX TISSUE=Peripheral blood;  
MEDLINE=93252853; PubMed=8486559;  
RA Farruti G., Peracchia F., Salliese M., Ambrosini G., Masini M.,  
RA Rotilio D., de Biasi A.;  
RT "Molecular analysis of human beta-arrestin-1: cloning, tissue  
distribution, and regulation of expression. Identification of two  
isoforms generated by alternative splicing.";  
RL J. Biol. Chem. 268:9753-9761 (1993).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOPROFORM 1A AND 1B).  
RC TISSUE=Brain;  
RX Yu Q.M., Zhou T.H., Cheng Z.J., Ma L., Pei G.;  
RT "Molecular cloning of two isoforms of human beta-arrestin 1.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOPROFORM 1A).  
RC TISSUE=Uterus;  
MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Haieh F.,  
RA Blatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Faney J., Helton E., Kettner M., Madan A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -1- FUNCTION: Regulates beta-adrenergic receptor function. Beta-  
arrestins seem to bind phosphorylated beta-adrenergic receptors,  
thereby causing a significant impairment of their capacity to  
activate G(s) proteins.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1A;  
CC IsoId=P49407-1; Sequences=Displayed;  
CC Name=1B;  
CC IsoId=P49407-2; Sequence=VSP\_000322;  
CC -1- SIMILARITY: Belongs to the arrestin family.  
CC



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L04685; AAA35558.1; --  
 CC EMBL; L04685; AAA35558.1; --  
 CC EMBL; AF084040; AAC33295.1; --  
 CC EMBL; AF084940; AAC34123.1; --  
 CC EMBL; BC003636; AAH03636.1; --  
 CC PIR; B45682; B45682.  
 CC HSSP; P08168; ICF1.  
 CC Genew; HGNC:711; ARRB1.  
 CC MIM; 107940; --  
 CC GO; GO:0005737; C:cytoplasm; TAS.  
 CC GO; GO:0005834; C:heterotrimeric G-protein complex; TAS.  
 CC GO; GO:0005886; C:plasma membrane; TAS.  
 CC GO; GO:0005625; C:soluble fraction; TAS.  
 CC GO; GO:0004857; F:enzyme inhibitor activity; TAS.  
 CC GO; GO:0007165; P:signal transduction; TAS.  
 CC InterPro; IPR000698; Arrestin.  
 CC InterPro; IPR007110; Ig-like.  
 CC Pfam; PF00339; arrestin; 1.  
 CC Pfam; PF02752; arrestin; 1.  
 CC PRINTS; PR00309; ARRESTIN.  
 CC ProDom; PD02099; Arrestin; 1.  
 CC PROSITE; PS00295; ARRESTINS; 1.  
 CC Sensory transduction; Alternative splicing.  
 CC VARSPLIC 334 341 Missing (in isoform 1B).  
 CC FT CONFLICT 146 146 V -> A (IN REF. 1).  
 CC FT CONFLICT 165 165 R -> G (IN REF. 1).  
 CC FT CONFLICT 229 229 K -> E (IN REF. 1).  
 CC FT CONFLICT 329 329 V -> E (IN REF. 2).  
 CC FT CONFLICT 400 400 K -> E (IN REF. 1).  
 CC FT CONFLICT 414 414 Q -> R (IN REF. 1).  
 CC FT CONFLICT 417 417 N -> D (IN REF. 1).  
 CC SQ SEQUENCE 418 AA; 47065 MW; 0A3C3C35092338D10 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 418;  
 Best Local Similarity 50.0%; Pred. No. 81;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RAQKVQLKLGK 12  
 Db 99 RLQERLIKLGK 110

RESULT 36  
 ARRI RAT ID ARRI RAT STANDARD; PRT; 418 AA.  
 AC P29066;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Beta-arrestin 1 (Arrestin, beta 1).  
 GN ARRB1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=92388146; PubMed=1517224;  
 RA Attamadal H., Arriza J.L., Aoki C., Dawson T.M., Codina J.,  
 RA Kwatra M.M., Snyder S.H., Caron M.G., Lefkowitz R.J.;  
 RA "Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene  
 RT family."; Chem. 267:17882-17890 (1992).  
 RT J. Biol.

CC -!- FUNCTION: Regulates beta-adrenergic receptor function. Beta-  
 CC arrestins seem to bind phosphorylated beta-adrenergic receptors,  
 CC thereby causing a significant impairment of their capacity to  
 CC activate G(S) proteins.  
 CC -!- TISSUE SPECIFICITY: Predominantly localized in neuronal tissues  
 CC and in the spleen.  
 CC -!- SIMILARITY: Belongs to the arrestin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M91589; AA74459.1; --  
 CC PIR; B43404; B43404.  
 CC HSSP; P08168; ICF1.  
 CC InterPro; IPR000698; Arrestin.  
 CC InterPro; IPR007110; Ig-like.  
 CC Pfam; PF00339; arrestin; 1.  
 CC Pfam; PF02752; arrestin; 1.  
 CC PRINTS; PR00309; ARRESTIN.  
 CC ProDom; PD02099; Arrestin; 1.  
 CC PROSITE; PS00295; ARRESTINS; 1.  
 CC Sensory transduction.  
 CC SQ SEQUENCE 418 AA; 47019 MW; 0A3C07D71B7ABC55 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 418;  
 Best Local Similarity 50.0%; Pred. No. 81;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RAQKVQLKLGK 12  
 Db 99 RLQERLIKLGK 110

RESULT 37  
 HEXA FORGI ID HEXA FORGI STANDARD; PRT; 777 AA.  
 AC P49008;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Beta-hexosaminidase precursor (EC 3.2.1.52) (N-acetyl-beta-  
 DE Glucosaminidase) (Beta-GlcNAcase) (Beta-N-acetylhexosaminidase)  
 DE (Beta-NAHase).  
 GN NAHA OR PG0043.  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Porphyromonadaceae; Porphyromonas.  
 OC NCBI\_TaxID=837;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=W83;  
 RX MEDLINE=95187310; PubMed=7881557;  
 RA Lovatt A., Roberts I.S.;  
 RA "Cloning and expression in Escherichia coli of the nahA gene from  
 RT Porphyromonas gingivalis indicates that beta-N-acetylhexosaminidase  
 RT is an outer-membrane-associated lipoprotein.";  
 RL Microbiology 140:3399-3406 (1994).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=W83;  
 RX MEDLINE=22829867; PubMed=12949112;  
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,  
 RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.,  
 RA Hatt D.H., Kohnen J.F., Nelson W.C., Mason T., Tallon L., Gray J.,  
 RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,  
 RA Dewhirst F.E., Fraser C.M.;  
 RA "Complete genome sequence of the oral pathogenic bacterium  
 RT Porphyromonas gingivalis strain W83.";

RL J. Bacteriol. 185:5591-5601 (2003).

CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-

CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.

CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid

CC anchor (Probable).

CC -1- SIMILARITY: Belongs to family 20 of glycosyl hydrolases.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; X78979; CAA55582.1; --

DR EMBL; AE011712; AAQ65295.1; --

DR TIGR; PG0043; --

DR InterPro; IPR001540; Glyco\_hydro\_20.

DR InterPro; IPR000437; Prok\_Lipopr\_t\_s.

DR Pfam; PF00728; Glyco\_hydro\_20; 1.

DR Pfam; PF02838; Glyco\_hydro\_20b; 1.

DR PRINTS; PR00738; GLHYDRLASE20.

DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.

DR Hydrolase; Glycosidase; Signal; Outer membrane; Lipoprotein;

KW Palmitate; Complete proteome.

FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 777 BETA-HEXOSAMINIDASE.

FT LIPID 19 19 N-palmitoyl cysteine (Probable).

FT LIPID 19 19 S-diacylglycerol cysteine (Probable).

FT CONFLICT 258 258 R -> H (IN REF. 1).

FT CONFLICT 265 265 E -> M (IN REF. 1).

FT CONFLICT 282 283 LA -> FR (IN REF. 1).

FT CONFLICT 575 575 T -> S (IN REF. 1).

FT CONFLICT 747 747 G -> A (IN REF. 1).

FT CONFLICT 777 777 AA; 87661 MW; DOA55D2C2FFAD864 CRC64;

SQ SEQUENCE 777 AA; 87661 MW; DOA55D2C2FFAD864 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 777;

Best Local Similarity 88.9%; Pred. No. 1.4e+02;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EKVLQKLGK 12

Db 371 EKVLQKRGK 379

RESULT 38

CHD3\_CABEL STANDARD; PRT; 1787 AA.

ID -CHD3\_CABEL STANDARD; PRT; 1787 AA.

AC Q22516; Q18794; 62

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chromodomain helicase-DNA-binding protein 3 homolog (CHD-3).

GN CHD-3 OR T14G8.1

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20530482; PubMed=11076750;

RA von Zelewsky T.; Palladino F.; Brunschwig K.; Tobler H.; Hajnal A.;

RA Mueller F.;

RT "The C. elegans Mi-2 chromatin-remodelling proteins function in vulval

RT cell fate determination."

RL Development 127:5277-5284 (2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Matthews P.; McMurray A.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Chromatin-remodelling protein that function in vulval

cell fate determination.

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.

CC -1- SIMILARITY: Contains 2 PHD-type zinc fingers.

CC -1- SIMILARITY: Contains 2 chromo domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; AF308444; RAQ29837.1; --

DR EMBL; Z67884; CAA31810.1; --

DR EMBL; Z67881; CAA31810.1; JOINED.

DR EMBL; Z67881; CAA31798.1; --

DR EMBL; Z67884; CAA31798.1; JOINED.

DR PIR; T20160; T20160.

DR WormPep; T14G8.1; CE03657.

DR InterPro; IPR000953; Chromo.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR002464; DEAH\_box.

DR InterPro; IPR001650; Helicase\_C.

DR InterPro; IPR000330; SNF2\_N.

DR InterPro; IPR001965; Znf\_FHD.

DR InterPro; IPR001841; Znf\_ring.

DR Pfam; PF00385; Chromo; 1.

DR Pfam; PF00271; Helicase\_C; 1.

DR Pfam; PF00628; PHD; 2.

DR Pfam; PF00176; SNF2\_N; 1.

DR SMART; SM00298; CHROMO; 2.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC; 1.

DR SMART; SM00249; PHD; 2.

DR SMART; SM00184; RING; 2.

DR PROSITE; PS00538; CHROMO\_1; FALSE\_NEG.

DR PROSITE; PS00013; CHROMO\_2; 1.

DR PROSITE; PS00690; DEAH\_ATP\_HELICASE; 1.

DR PROSITE; PS01359; ZF\_PHD\_1\_2.

DR PROSITE; PS00016; ZF\_PHD\_2\_2.

KW Chromatin regulator; Nuclear protein; Repeat; Helicase; DNA-binding;

KW ATP-binding; Zinc-finger.

FT DOMAIN 59 62 POLY-LYS.

FT ZN\_FING 265 312 PHD-TYPE 1.

FT ZN\_FING 328 375 PHD-TYPE 2.

FT DOMAIN 373 476 CHROMO 1.

FT DOMAIN 501 583 CHROMO 2.

FT DOMAIN 1287 1291 POLY-ARG.

FT NP\_BIND 641 648 ATP (POTENTIAL).

FT SITE 763 766 DEAH\_BOX.

SQ SEQUENCE 1787 AA; 205254 MW; 1EFCB1FPECB59740 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 1787;

Best Local Similarity 80.0%; Pred. No. 2.9e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QEKVLQKLGK 12

Db 1226 QETELQKLGK 1235

RESULT 39

STAR\_PIG STANDARD; PRT; 285 AA.

ID -STAR\_PIG STANDARD; PRT; 285 AA.

AC Q28996; Q95259;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Steroidogenic acute regulatory protein, mitochondrial precursor

DE (STAR) (STARD1).

GN STAR.

OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Ovary;  
RX MEDLINE=97200784; PubMed=9048613;  
RA Pilon N., Daneau I., Brissou C., Ethier J.P., Lussier J.G.,  
RA Silversides D.W.;  
RT "Porcine and bovine steroidogenic acute regulatory protein (STAR)  
RT gene expression during gestation.";  
RL Endocrinology 138:1085-1091(1997).  
RN [2]  
RP SEQUENCE OF 132-210 FROM N.A.  
RX TISSUE=Corpus luteum;  
RX MEDLINE=97131976; PubMed=8977433;  
RA Balasubramanian K., Lavoie H.A., Garney J.C., Stocco D.M.,  
RA Velhous J.D.;  
RT "Regulation of porcine granulosa cell steroidogenic acute regulatory  
RT protein (STAR) by insulin-like growth factor I: synergism with  
RT follicle-stimulating hormone or protein kinase A agonist.";  
RL Endocrinology 138:433-439(1997).  
CC -!- FUNCTION: Plays a key role in steroid hormone synthesis by  
CC enhancing the metabolism of cholesterol into pregnenolone.  
CC Mediates the transfer of cholesterol from the outer mitochondrial  
CC membrane to the inner mitochondrial membrane where it is cleaved  
CC to pregnenolone (By similarity).  
CC -!- PATHWAY: Steroidogenesis; rate-limiting step.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
CC -!- TISSUE SPECIFICITY: Expressed in adult and fetal steroidogenic  
CC tissues, including adult testes and ovaries and adult adrenal  
CC glands as well as steroidogenic tissues of pregnancy, including  
CC developing fetal testes, corpus luteum, and pregnancy, but not in  
CC the fetal ovary.  
CC -!- SIMILARITY: Contains 1 START domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U53020; AAB04953.1; -.  
CC EMBL; U72195; AAB50555.1; -.  
CC HS67; Q14849; IEM2.  
CC InterPro; IPR000799; STAR.  
CC Pfam; PF01852; START; 1.  
CC PRINTS; PR00978; STARPROTEIN.  
CC SMART; SM00234; START; 1.  
CC PROSITE; PS00848; START; 1.  
CC Lipid-binding; Lipid transport; Transport; Steroidogenesis;  
CC Mitochondrion; Transit peptide; Phosphorylation.  
CC TRANSIT 1 MITOCHONDRION (POTENTIAL).  
CC CHAIN ? 285 STEROIDGENIC ACUTE REGULATORY PROTEIN.  
CC DOMAIN 67 280 START.  
CC MOD RES 57 57 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
CC MOD RES 195 195 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
CC CONFLICT 188 188 G -> R (IN REF. 2).  
CC SEQUENCE 285 AA; 31833 MW; 67DCA9DA521519B8 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 285;  
Best Local Similarity 75.0%; Pred. No. 86;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KVLQKLGK 12  
Db 155 KILQKIGK 162

RESULT 40  
STAR\_XENLA  
ID STAR\_XENLA STANDARD; PRT; 289 AA.  
AC Q9DQ08;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Steroidogenic acute regulatory protein (STAR) (STARD1) (Fragment).  
GN STAR.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodidae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Ovary;  
RX MEDLINE=20519388; PubMed=11064158;  
RA Bauer M.P., Bridgham J.T., Langenau D.M., Johnson A.L., Goetz F.W.;  
RT "Conservation of steroidogenic acute regulatory (STAR) protein  
RT structure and expression in vertebrates.";  
RL Mol. Cell. Endocrinol. 168:119-125(2000).  
CC -!- FUNCTION: Plays a key role in steroid hormone synthesis by  
CC enhancing the metabolism of cholesterol into pregnenolone.  
CC Mediates the transfer of cholesterol from the outer mitochondrial  
CC membrane to the inner mitochondrial membrane where it is cleaved  
CC to pregnenolone (By similarity).  
CC -!- PATHWAY: Steroidogenesis; rate-limiting step.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
CC -!- SIMILARITY: Contains 1 START domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF220437; AAG28595.1; -.  
CC HS67; Q14849; IEM2.  
CC InterPro; IPR000799; STAR.  
CC InterPro; IPR002913; START.  
CC Pfam; PF01852; START; 1.  
CC PRINTS; PR00978; STARPROTEIN.  
CC SMART; SM00234; START; 1.  
CC PROSITE; PS00848; START; 1.  
CC Lipid-binding; Lipid transport; Transport; Steroidogenesis.  
CC NON TER 1  
CC DOMAIN 71 284 START.  
CC NON TER 289 289  
CC SEQUENCE 289 AA; 32066 MW; 0772112AB886BD8E CRC64;  
Query Match 57.4%; Score 35; DB 1; Length 289;  
Best Local Similarity 75.0%; Pred. No. 87;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 5 KVLQKLGK 12  
Db 159 KILQKIGK 166

RESULT 41  
PARB\_CAUCR  
ID PARB\_CAUCR STANDARD; PRT; 294 AA.  
AC O05190;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Chromosome partitioning protein parB.  
GN PARB OR CC3752.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;

CC Caulobacteraceae; Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CB15N / NA1000;  
RX MEDLINE=97207014; PubMed=9054507;  
RA Mohl D.A., Guber J.W.;  
RT "Cell cycle-dependent polar localization of chromosome partitioning  
proteins in Caulobacter crescentus";  
RN Cell 88:675-684(1997).  
CC [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
R Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
R Potocka I., Nelson W.C., Newton A.S., Stephens C., Haft D.H.,  
R DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
R Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
R Utterback T., Tran K., Wolf A., Vamathevan J., Ertolaeva M., White O.,  
R Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RA "Complete genome sequence of Caulobacter crescentus";  
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
CC -!- FUNCTION: INVOLVED IN CHROMOSOME PARTITION. LOCALIZE TO BOTH POLES  
OF THE PREDIVISIONAL CELL FOLLOWING COMPLETION OF DNA REPLICATION.  
CC BINDS TO THE DNA ORIGIN OF REPLICATION.  
CC -!- SIMILARITY: Belongs to the parB family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL; U87804; AAB51268.1; -;  
DR EMBL; AE006032; AAK25714.1; -;  
DR PIR; F87714; F87714.  
DR TIGR; CC3752; -;  
DR InterPro; IPR004437; ParB\_part.  
DR InterPro; IPR003115; ParBC.  
DR Pfam; PF02195; ParBC; 1.  
DR SMART; SM00470; ParB; 1.  
DR TIGRFAMs; TIGR00180; parB part; 1.  
KW Chromosome partition; DNA-binding; Complete proteome.  
FT CONFLICT 15 26 LLGEVDAAPAAQ -> CWASRRRAGSG (IN REF. 1).  
FT CONFLICT 31 31 L -> F (IN REF. 1).  
FT CONFLICT 294 AA; 31824 MW; FAYD3DCP8A1AE639 CRC64;  
SQ SEQUENCE 294 AA; 31824 MW; 35686 MW; 6DF4703EA7252910 CRC64;  
Query Match 57.4%; Score 35; DB 1; Length 294;  
Best Local Similarity 46.2%; Pred. No. 88;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 1 RAQEKVLQKLGKA 13  
DB 150 RTQENIAQTIGKS 162  
RESULT 42  
Y006 YEAST  
ID Y006 YEAST STANDARD; PRT; 313 AA.  
AC Q12094;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Hypothetical UPF0293 protein YOR006C.  
GN YOR006C OR UND313.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=97051599; PubMed=8896276;  
RA Sterky F., Holmberg A., Pettersson B., Uhlen M.;  
RT "The sequence of a 30 kb fragment on the left arm of chromosome XV  
from Saccharomyces cerevisiae reveals 15 open reading frames, five of  
which correspond to previously identified genes";  
RL Yeast 12:1091-1095(1996).  
CC -!- SIMILARITY: Belongs to the UPF0293 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL; U43491; AAC49486.1; -;  
DR EMBL; Z74914; CAA99194.1; -;  
DR PIR; S61990; S61990.  
DR GerMOnline; 143594; -;  
DR SGD; S0005532; YOR006C.  
DR InterPro; IPR007177; DUF367.  
DR InterPro; IPR007209; RLI.  
DR Pfam; PF04034; DUF367; 1.  
DR Pfam; PF04068; RLI; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 313 AA; 35686 MW; 6DF4703EA7252910 CRC64;  
Query Match 57.4%; Score 35; DB 1; Length 313;  
Best Local Similarity 66.7%; Pred. No. 93;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RAQEKVLQKLGK 12  
DB 204 RAEEWLOKLEK 215  
RESULT 43  
AR22 HUMAN  
ID AR22 HUMAN STANDARD; PRT; 409 AA.  
AC P32121; O8N7Y2; Q9UEQ6;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Beta-arrestin 2 (Arrestin, beta 2).  
GN AR22 OR AR22 OR AR2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Thyroid;  
RX MEDLINE=92267237; PubMed=1587386;  
RA Rapoport B., Kaufman K.D., Chamenbalk G.D.;  
RT "Cloning of a member of the arrestin family from a human thyroid cDNA  
library";  
RL Mol. Cell. Endocrinol. 84:R39-R43(1992).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Brain;  
RA Yu Q.M., Zhou T.H., Wu Y.L., Cheng Z.J., Ma L., Pei G.;  
RT "G-protein coupled receptor interaction with beta-arrestin 2 through  
specific agonist stimulation";  
RN Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
CC [3]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Testis;  
RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,  
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,  
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,  
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,

RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,  
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,  
RA Sugano S., Nagahara K., Masuho Y., Nagai K., Isegai T.;  
RT "NEDO human cDNA sequencing project";  
RN Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
[4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RL [1]  
CC -!- FUNCTION: Regulates beta-adrenergic receptor function. Beta-  
arrestins seem to bind phosphorylated beta-adrenergic receptors,  
thereby causing a significant impairment of their capacity to  
activate G(s) proteins.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Name=1;  
CC Name=2;  
CC IsoId=P32121-1; Sequence=Displayed;  
CC IsoId=P32121-2; Sequence=VSP\_008194, VSP\_008195;  
CC Note=No experimental confirmation available;  
CC -!- SIMILARITY: Belongs to the arrestin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z11501; CAA77577.1; -;  
CC EMBL; AF106941; AAC99468.1; -;  
CC EMBL; AK097542; BAC05094.1; -;  
CC EMBL; BC007427; AAH07427.1; -;  
CC F1R; S18984; S18984.  
CC HSP; P08168; 1CPI.  
CC Genew; HGNC:712; ARRB2.  
CC MIM; 107941; -;  
CC GO; GO:0005737; C:cytoplasm; TAS.  
CC GO; GO:0005886; C:plasma membrane; TAS.  
CC InterPro; IPR000698; Arrestin.  
CC InterPro; IPR007110; IG-like.  
CC Pfam; PF00339; arrestin; 1.  
CC Pfam; PF02752; arrestin; C; 1.  
CC PRINTS; P00309; ARRESTIN.  
CC ProDom; PD002059; Arrestin; 1.  
CC PROSITE; PS00295; ARRESTINS; 1.  
KW Sensory transduction; Nuclear protein; Alternative splicing.  
FT DOMAIN 39 399 NUCLEAR EXPORT SIGNAL (By similarity).  
FT VARSPLIC 39 53 Missing (in isoform 2).  
FT FTId=VSP\_008194.  
FT S -> SAPTPPTPLPVPP (in isoform 2).  
FT

FT  
FT CONFLICT 189 189 /FTId=VSP\_008195.  
FT R -> P (IN REF. 1).  
SQ SEQUENCE 409 AA; 46105 MW; DSEC507DA47B84FF CRC64;  
Query Match 57.4%; Score 35; DB 1; Length 409;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RAQKVLQKLGK 12  
Db 100 RLQDRLLRLQK 111  
RESULT 44  
ARR2 MOUSE  
ID ASR2 MOUSE STANDARD; PRT; 410 AA.  
AC QSLVTA;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Beta-arrestin 2 (Arrestin, beta 2).  
GN ARRB2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RL [2]  
RP NUCLEOCYTOPLASMIC SHUTTLING.  
RX MEDLINE=22241956; PubMed=12167659;  
RA Scott M.G., Le Rouzic E., Perianin A., Pierotti V., Enslin H.,  
RA Benichou S., Marullo S., Benmerah A.;  
RT "Differential nucleocytoplasmic shuttling of beta-arrestins.  
Characterization of a leucine-rich nuclear export signal in  
beta-arrestin2";  
RL J. Biol. Chem. 277:37693-37701(2002).  
CC -!- FUNCTION: Regulates beta-adrenergic receptor function. Beta-  
arrestins seem to bind phosphorylated beta-adrenergic receptors,  
thereby causing a significant impairment of their capacity to  
activate G(s) proteins.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.  
CC -!- TISSUE SPECIFICITY: Predominantly localized in neuronal tissues  
and in the spleen.  
CC -!- SIMILARITY: Belongs to the arrestin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; BC016642; AAH16642.1; -  
MGD; MGI:99474; Artd2.  
InterPro; IPR000698; Arrestin.  
InterPro; IPR007110; Ig-like.  
Pfam; PF00339; arrestin; 1.  
Pfam; PF02752; arrestin C; 1.  
PRINTS; PR00309; ARRESTIN.  
PRODOM; PD002099; Arrestin; 1.  
PROSITE; PS00295; ARRESTINS; 1.  
Sensory transduction; Nuclear protein.  
FT DOMAIN 390 400 NUCLEAR EXPORT SIGNAL.  
SQ SEQUENCE 410 AA; 46314 MW; ODFA73A1C532AE03 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 410;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
| : : : : :  
Db 100 RLQDRLLKLGK 111

RESULT 45  
ID ARR2 RAT STANDARD; PRT; 410 AA.  
AC P29067;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Beta-arrestin 2 (Arrestin, beta 2).  
GN ARRB2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=92388146; PubMed=1517224;  
RA Attramadali H., Arriza J.L., Acki C., Dawson T.M., Codina J.,  
RA Kwatra M.M., Snyder S.H., Caron M.G., Lefkowitz R.J.;  
RT "Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene  
RT family.";  
RL J. Biol. Chem. 267:17882-17890(1992).  
RN [2]  
RP SEQUENCE OF 295-410 FROM N.A.  
RC TISSUE=Pineal gland;  
RX MEDLINE=94140898; PubMed=8308033;  
RA Craft C.M., Whitmore D.H., Wlechniak A.F.;  
RT "Cone arrestin identified by targeting expression of a functional  
RT family.";  
RL J. Biol. Chem. 269:4613-4619(1994).  
RN [3]  
RP SEQUENCE OF 305-386 FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Retina;  
RX MEDLINE=99145674; PubMed=9767391;  
RA Komori N., Cain S.D., Roch J.-M., Miller K.E., Matsumoto H.;  
RT "Differential expression of alternative splice variants of  
RT beta-arrestin-1 and -2 in rat central nervous system and peripheral  
RT tissues.";  
RL Eur. J. Neurosci. 10:2607-2616(1998).  
CC -!- FUNCTION: Regulates beta-adrenergic receptor function. Beta-  
CC arrestins seem to bind phosphorylated beta-adrenergic receptors,  
CC thereby causing a significant impairment of their capacity to  
CC activate G(s) proteins.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
CC -!- TISSUE SPECIFICITY: Predominantly localized in neuronal tissues  
CC and in the spleen.  
CC -!- SIMILARITY: Belongs to the arrestin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M91590; AAA74460.1; -  
EMBL; U03627; AAA17551.1; -  
EMBL; AF051457; AAC28617.1; -  
PIR; A59279; A59279.  
HSP; P08168; ICPI.  
InterPro; IPR000698; Arrestin.  
InterPro; IPR007110; Ig-like.  
Pfam; PF00339; arrestin; 1.  
Pfam; PF02752; arrestin C; 1.  
PRINTS; PR00309; ARRESTIN.  
PRODOM; PD002099; Arrestin; 1.  
PROSITE; PS00295; ARRESTINS; 1.  
Sensory transduction; Nuclear protein.  
FT DOMAIN 390 400 NUCLEAR EXPORT SIGNAL (BY SIMILARITY).  
SQ SEQUENCE 410 AA; 46340 MW; ODFA6A897C2B86BA CRC64;

Query Match 57.4%; Score 35; DB 1; Length 410;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGK 12  
| : : : : :  
Db 100 RLQDRLLKLGK 111

RESULT 46  
ID ARLY OCEIH STANDARD; PRT; 459 AA.  
AC OBLT9;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Argininosuccinate lyase (EC 4.3.2.1) (Argininosuccinase) (ASAL).  
GN ARGH OR OB3128.  
OS Oceanobacillus iheyensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;  
RX MEDLINE=2220767; PubMed=12235376;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya  
RT Ridge and its unexpected adaptive capabilities to extreme  
RT environments.";  
RL Nucleic Acids Res. 30:3927-3935(2002).  
CC -!- CATALYTIC ACTIVITY: N-(L-arginino)succinate = fumarate + L-  
CC arginine.  
CC -!- PATHWAY: Arginine biosynthesis; eighth (last) step.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: Belongs to the lyase 1 family. Argininosuccinate lyase  
CC subfamily.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AP004603; BAC15084.1; -  
HMAP; MF 00006; -; 1.  
InterPro; IPR005049; argH.  
InterPro; IPR000362; Fumarate\_lyase.  
InterPro; IPR008948; L-Aspartase-like.

```
DR Pfam: PF00206; lyase 1; 1.
DR PRINTS: PR00149; FUMATELYASE.
DR TIGRFAMs: TIGR00838; argH; 1.
DR PROSITE: PS00163; FUMARATE LYASES; 1.
DR Arginine biosynthesis; Lyase; Complete proteome.
KW SEQUENCE 459 AA; 51230 MW; C100878/5B890964 CRC64;
SQ
Query Match 57.4%; Score 35; DB 1; Length 459;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 AOEKVLQKLGKA 13
|:|:|:|:|
DB 440 AKESVLQOIAXA 451

RESULT 47
GATE_STRCO
ID GATE_STRCO STANDARD; PRT; 504 AA.
AC Q92578;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Asparyl-glutamyl-tRNA(Asn/Gln) amidotransferase subunit B
DE (EC 6.3.5.-) (Asp/Glu-ADT subunit B).
GN GATB OR SC05501 OR SC899.13.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble L., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: Allows the formation of correctly charged Asn-tRNA(Asn)
CC or Gln-tRNA(Gln) through the transamidation of misacylated Asp-
CC tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both
CC of asparaginyl-tRNA or glutamyl-tRNA synthetases. The reaction
CC takes place in the presence of glutamine and ATP through an
CC activated phospho-Asp-tRNA(Asn) or phospho-Glu-tRNA(Gln) (by
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Gln) + L-glutamine = ADP
CC + phosphate + L-glutamyl-tRNA(Gln) + L-glutamate.
CC -!- CATALYTIC ACTIVITY: ATP + L-asparaginyl-tRNA(Asn) + L-glutamine = ADP
CC + phosphate + L-asparaginyl-tRNA(Asn) + L-glutamate.
CC -!- SUBUNIT: Heterotrimer of A, B and C subunits (By similarity).
CC -!- SIMILARITY: Belongs to the gatB/gatE family. GatB subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL939124; CAB37577.1; -.
CC PIR; T35817; T35817.
CC HAMAP; MF_00121; -.
CC InterPro; IPR004413; GatB.
CC InterPro; IPR006107; GatB_cent.
CC

DR Pfam: PF00206; lyase 1; 1.
DR PRINTS: PR00149; FUMATELYASE.
DR TIGRFAMs: TIGR00838; argH; 1.
DR PROSITE: PS00163; FUMARATE LYASES; 1.
DR Arginine biosynthesis; Lyase; Complete proteome.
KW SEQUENCE 459 AA; 51230 MW; C100878/5B890964 CRC64;
SQ
Query Match 57.4%; Score 35; DB 1; Length 459;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 AOEKVLQKLGKA 13
|:|:|:|:|
DB 440 AKESVLQOIAXA 451

RESULT 48
REC_N_HELPJ
ID REC_N_HELPJ STANDARD; PRT; 522 AA.
AC Q9ZJ80;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein recN (Recombination protein N).
DE REC_N_HELPJ.
GN REC_N_HELPJ.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.B., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RA "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
CC -!- DNA (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the recN family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001566; AB007019.1; -.
CC PIR; B71807; B71807.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC SMART; SM00382; AAA; 1.
CC NP_BIND 36 43 ATP (POTENTIAL).
FT NP_BIND 36 43
SQ SEQUENCE 522 AA; 59349 MW; 000D7EE3C0632EAS CRC64;
QY 4 EKVLEKLG 11
|:|:|:|
DB 282 EKVLEKLG 289

Query Match 57.4%; Score 35; DB 1; Length 522;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 49  
 NAF1 HUMAN STANDARD; PRT; 636 AA.  
 ID Q15025; Q76008; Q96EL9; Q99833; Q9H1J3;  
 AC 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Nef-associated factor 1 (Naf1) (Hiv-1 Nef interacting protein)  
 DE (Viron-associated nuclear shuttling protein) (VAN) (hVAN) (TNFAIP3  
 DE interacting protein 1)  
 GN TNFIP1 OR NAF1 OR KIAA0113.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Peripheral blood;  
 RX MEDLINE=99120485; PubMed=9923610;  
 RA Fukushi M., Dixon J., Kimura T., Teurutani N., Dixon M.J.,  
 Yamamoto N.;  
 RT "Identification and cloning of a novel cellular protein Naf1, Nef-  
 RT associated factor 1, that increases cell surface CD4 expression.";  
 RL FEBS Lett. 442:83-88 (1999).  
 [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=20541981; PubMed=11090181;  
 RA Gupta K., Ott D., Hope T.J., Siliciano R.P., Boeke J.D.;  
 RT "A human nuclear shuttling protein that interacts with human  
 RT immunodeficiency virus type 1 matrix is packaged into virions.";  
 RL J. Virol. 74:11811-11824 (2000).  
 [3]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Lung;  
 RX MEDLINE=22386257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenner C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,  
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [4]  
 RN SEQUENCE OF 136-636 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Craniofacial;  
 RX MEDLINE=96276047; PubMed=8681136;  
 RA Loftus S.K., Dixon J., Koprivnikar K., Dixon M.J., Wasmuth J.J.;  
 RT "Transcriptional map of the Treacher Collins candidate gene region.";  
 RL Genome Res. 6:26-34 (1996).  
 [5]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=95308325; PubMed=7788527;  
 RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,  
 RA Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;  
 RN Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 [6]

RT "Prediction of the coding sequences of unidentified human genes. III.  
 RT The coding sequences of 40 new genes (K1AA0081-K1AA0120) deduced by  
 RL analysis of cDNA clones from human cell line KG-1.";  
 RL DNA Res. 2:37-43 (1995).  
 [7]  
 RN SEQUENCE OF 94-412 FROM N.A.  
 RA Fukushi M., Kimura T., Yamamoto N.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -P- FUNCTION: Interacts with zinc finger protein A20/TNFAIP3 and  
 CC inhibits TNF-induced NF-kappa-B-dependent gene expression by  
 CC interfering with an RIP- or TRAF2-mediated transactivation signal  
 CC (By similarity). Increases cell surface CD4 (T4) antigen  
 CC expression. Interacts with HIV-1 matrix protein and is packaged  
 CC into virions and overexpression can inhibit viral replication. May  
 CC regulate matrix nuclear localization, both nuclear import of PIC  
 CC (preintegration complex) and export of GAG polypeptide and viral  
 CC genomic RNA during virion production.  
 CC -I- SUBUNIT: Interacts with TNFAIP3 (By similarity). Interacts with  
 CC HIV-1 matrix protein.  
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic. Shuttles between the nucleus  
 CC and cytoplasm in a CRM1-dependent manner.  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=Alpha;  
 CC IsoId=Q15025-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Beta;  
 CC IsoId=Q15025-2; Sequence=VSP\_003913;  
 CC Note=No experimental confirmation available;  
 CC -I- TISSUE SPECIFICITY: Ubiquitous. Strongly expressed in peripheral  
 CC blood lymphocytes, spleen and skeletal muscle, and is weakly  
 CC expressed in the brain.  
 CC -I- CAUTION: Ref.7 sequence differs from that shown due to frameshifts  
 CC in positions 152 and 154.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ011895; CAA03855.1; -  
 DR EMBL; AJ011896; CAA03856.1; -  
 DR EMBL; AY012155; AAG42154.1; -  
 DR EMBL; BC012133; AAH12133.1; -  
 DR EMBL; BC014008; AAH14008.1; -  
 DR EMBL; U39403; AAC99999.1; -  
 DR EMBL; D30755; BAA06416.2; -  
 DR EMBL; U83844; AAB41438.1; ALT\_FRAME.  
 DR Genew; HGNC:16903; TNFIP1.  
 DR MIM; 607714; -  
 DR GO; GO:0005622; C:intracellular; TAS.  
 DR GO; GO:0005515; P:protein binding; TAS.  
 DR GO; GO:0009101; P:glycoprotein biosynthesis; IDA.  
 DR GO; GO:0045071; P:negative regulation of viral genome replica. .; TAS.  
 KW Coiled coil; Nuclear protein; Alternative splicing.  
 FT DOMAIN 20 73  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 196 258  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 294 535  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 94 412  
 FT INTERACTS WITH NEF.  
 FT DOMAIN 524 530  
 FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 539 636  
 FT PRO-RICH.  
 FT VARSPLIC 627 636  
 FT SPKNDREGPQ -> PADLRLPRN (in isoform 2).  
 FT /FTID=VSP\_003913.  
 FT CONFLICT 148 148  
 FT G -> D (IN REF. 3; AAH12133).  
 FT CONFLICT 299 299  
 FT A -> P (IN REF. 2).  
 FT SEQUENCE 636 AA; 71864 MW; D81B96BEAD50D871 CRC64;  
 SQ

Query Match 57.4%; Score 35; DB 1; Length 636;  
 Best Local Similarity 63.6%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;



```

CC Name-3;
CC isoid-Q9WU8-3; Sequence-VSP_003915;
CC TISSUE SPECIFICITY: Ubiquitous. Abundant in heart and skeletal
CC muscle and expressed at lower levels in thymus, liver, kidney,
CC brain and intestinal tract.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ242777; CAB44239.1; -
CC EMBL; AJ242778; CAB44240.1; -
CC EMBL; BC008186; AAH08186.1; -
CC EMBL; BC008665; AAH08665.1; -
CC EMBL; AX011241; CAC07546.1; -
CC EMBL; AY012159; AAG42155.2; -
CC MGB; MGI11926184; Th1p1.
CC GO; GO:0005737; Cytoplasm; IDA.
CC coiled coil; Nuclear protein; Alternative splicing.
CC DOMAIN 39 72 COILED COIL (POTENTIAL).
CC DOMAIN 209 270 COILED COIL (POTENTIAL).
CC FT DOMAIN 311 551 COILED COIL (POTENTIAL).
CC FT DOMAIN 95 425 INTERACTS WITH NEF.
CC FT DOMAIN 537 543 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 552 647 PRO-RICH.
CC FT DOMAIN 162 165 POLY-PRO.
CC FT VARSPLIC 1 53 Missing (in isoform 2).
CC FT FT VARSPLIC 638 647 /FTIG-VSP_003914.
CC VARSPLIC 638 647 SADNDCDGPQ -> PADLRLPKV (in isoform 3).
CC FT /FTID-VSP_003915.
CC FT VQ -> TR (IN REF. 3).
CC FT CONFLICT 97 98 A -> V (IN REF. 2; AAH08186).
CC FT CONFLICT 308 308 A -> V (IN REF. 2; AAH08186).
CC FT CONFLICT 533 533 A -> V (IN REF. 2; AAH08186).
CC FT CONFLICT 647 AA; 73050 MW; 4280879A8C24A16E CRC64;
CC SEQUENCE
CC
CC Query Match 57.4%; Score 35; DB 1; Length 647;
CC Best Local Similarity 63.8%; Pred. No. 1.e-02;
CC Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 3 QEKVLQKLGRKA 13
CC |||||:|:|
CC 425 QEKELQELNKA 435
CC
CC Search completed: March 4, 2004, 17:44:49
CC Job time : 12.8065 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 17:38:36 ; Search time 30.6129 Seconds  
(without alignments)

133.987 Million cell updates/sec

Title: US-10-069-540A-2\_COPY\_23\_35

Perfect score: 61

Sequence: 1 RAQEKVLQKLGA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mbc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	250	11 Q8BXH2	Q8BXH2 mus musculus
2	61	100.0	385	4 Q8NPL3	Q8NPL3 homo sapien
3	61	100.0	439	4 Q8BTH3	Q8BTH3 homo sapien
4	61	100.0	478	13 Q7ZNP5	Q7ZNP5 xenopus lae
5	61	100.0	490	4 Q8WHH9	Q8WHH9 homo sapien
6	61	100.0	564	4 Q9UKN4	Q9UKN4 homo sapien
7	61	100.0	565	4 Q9UBW5	Q9UBW5 homo sapien
8	61	100.0	565	4 Q86VV0	Q86VV0 homo sapien
9	61	100.0	686	11 Q7TQF7	Q7TQF7 mus musculus
10	61	100.0	695	4 Q8N4G0	Q8N4G0 homo sapien
11	57	93.4	524	13 Q803I3	Q803I3 brachydanio
12	49	80.3	414	5 Q8WQ54	Q8WQ54 geodia cydo
13	47	77.0	602	5 Q9YQ92	Q9YQ92 drosophila
14	41	67.2	461	10 Q9FP28	Q9FP28 oryza sativ
15	41	67.2	570	16 P94544	P94544 bacillus su
16	41	67.2	651	16 Q8E5M0	Q8E5M0 streptococc

17	41	67.2	651	16 Q8DX0	Q8DX0 streptococc
18	41	67.2	874	10 Q84PY5	Q84PY5 oryza sativ
19	41	67.2	876	10 Q943V8	Q943V8 oryza sativ
20	41	67.2	876	10 Q7XXG7	Q7XXG7 oryza sativ
21	41	67.2	1235	5 Q8SX64	Q8SX64 drosophila
22	41	67.2	1381	5 Q9VZT7	Q9VZT7 drosophila
23	40	65.6	176	16 Q9EML3	Q9EML3 deinoxococcus
24	39	63.9	502	16 Q8J132	Q8J132 tropheryma
25	39	63.9	505	16 Q97PES	Q97PES streptococc
26	39	63.9	510	16 Q8DNU7	Q8DNU7 streptococc
27	39	63.9	513	16 Q8RBN7	Q8RBN7 thermococ
28	39	63.9	516	10 Q42701	Q42701 catharanthu
29	39	63.9	522	16 Q8J332	Q8J332 tropheryma
30	39	63.9	524	10 Q42700	Q42700 catharanthu
31	39	63.9	700	16 Q8ET76	Q8ET76 corynebacte
32	38	62.3	88	4 Q9BS66	Q9BS66 homo sapien
33	38	62.3	193	4 Q8IU96	Q8IU96 homo sapien
34	38	62.3	250	4 Q9UEN2	Q9UEN2 homo sapien
35	38	62.3	257	4 Q8IX25	Q8IX25 homo sapien
36	38	62.3	300	17 Q8TT25	Q8TT25 methanosarc
37	38	62.3	331	4 Q9S278	Q9S278 homo sapien
38	38	62.3	408	13 Q7T2D2	Q7T2D2 brachydanio
39	38	62.3	445	16 Q89AC2	Q89AC2 buchnera ap
40	38	62.3	452	16 Q8FJR9	Q8FJR9 xanthomonas
41	38	62.3	461	5 Q21004	Q21004 caenorhabdi
42	38	62.3	474	16 Q9KTE0	Q9KTE0 vibrio chol
43	38	62.3	474	16 Q8RCI2	Q8RCI2 salmonella
44	38	62.3	474	16 Q8DFE8	Q8DFE8 vibrio vuln
45	38	62.3	474	16 Q8CX45	Q8CX45 shewanella
46	38	62.3	474	16 Q87RP4	Q87RP4 vibrio para
47	38	62.3	475	16 Q83LY3	Q83LY3 shigella fl
48	38	62.3	475	16 Q87AP4	Q87AP4 xylella fas
49	38	62.3	475	16 Q7VE74	Q7VE74 haemophilus
50	38	62.3	485	16 Q8E8B5	Q8E8B5 xanthomonas
51	38	62.3	488	16 Q8Z8G5	Q8Z8G5 salmonella
52	38	62.3	496	16 Q9AD76	Q9AD76 streptomyce
53	38	62.3	497	16 Q9PEX2	Q9PEX2 xylella fas
54	38	62.3	509	5 Q9VTZ7	Q9VTZ7 drosophila
55	38	62.3	535	17 Q8ZZQ9	Q8ZZQ9 pyrobaculum
56	38	62.3	552	4 Q9N089	Q9N089 homo sapien
57	38	62.3	552	11 Q91YNO	Q91YNO mus musculu
58	38	62.3	574	16 Q9ZLM0	Q9ZLM0 rhizobium m
59	38	62.3	817	16 Q83DJ0	Q83DJ0 coxiella bu
60	38	62.3	908	17 Q8Q0E6	Q8Q0E6 methanosarc
61	37	60.7	81	2 Q34794	Q34794 methanobact
62	37	60.7	81	2 Q03115	Q03115 streptococc
63	37	60.7	168	2 Q9AGA2	Q9AGA2 streptococc
64	37	60.7	196	16 Q7VLQ5	Q7VLQ5 haemophilus
65	37	60.7	316	11 Q8BHS4	Q8BHS4 mus musculu
66	37	60.7	386	17 Q97C57	Q97C57 thermoplas
67	37	60.7	386	17 Q97A29	Q97A29 thermoplas
68	37	60.7	393	17 Q979E8	Q979E8 thermoplas
69	37	60.7	395	10 Q7XQW1	Q7XQW1 oryza sativ
70	37	60.7	403	17 Q26176	Q26176 methanobact
71	37	60.7	441	5 Q8MP09	Q8MP09 caenorhabdi
72	37	60.7	444	11 Q8JZT3	Q8JZT3 mus musculu
73	37	60.7	461	4 Q8N2L3	Q8N2L3 homo sapien
74	37	60.7	461	4 Q8H173	Q8H173 homo sapien
75	37	60.7	465	11 Q9EPK6	Q9EPK6 mus musculu
76	37	60.7	465	11 Q91V34	Q91V34 mus musculu
77	37	60.7	468	10 Q94J76	Q94J76 oryza sativ
78	37	60.7	472	10 Q98RQ7	Q98RQ7 guillardia
79	37	60.7	526	17 Q98616	Q98616 pyrococcus
80	37	60.7	545	16 Q8EEX1	Q8EEX1 shewanella
81	37	60.7	585	11 Q8BH7	Q8BH7 mus musculu
82	37	60.7	591	10 Q9LU46	Q9LU46 arabidopsis
83	37	60.7	636	10 Q7XWT1	Q7XWT1 oryza sativ
84	37	60.7	669	10 Q7XRD9	Q7XRD9 oryza sativ
85	37	60.7	684	10 Q7X8V4	Q7X8V4 oryza sativ
86	37	60.7	728	10 Q7XWRO	Q7XWRO oryza sativ
87	37	60.7	784	16 Q88QC2	Q88QC2 pseudomonas
88	37	60.7	820	10 Q7XLC2	Q7XLC2 oryza sativ
89	37	60.7	823	10 Q7XPA8	Q7XPA8 oryza sativ

90 37 60.7 839 10 Q7XP88  
 91 37 60.7 858 10 Q7XMB9  
 92 37 60.7 876 10 Q7XV39  
 93 37 60.7 876 10 Q7XNA9  
 94 37 60.7 897 10 Q8S6T5  
 95 37 60.7 905 11 Q8CHD8  
 96 37 60.7 1110 3 Q8NIU9  
 97 37 60.7 1148 10 Q8H897  
 98 37 60.7 1209 4 Q7Z570  
 99 37 60.7 1274 10 Q9ZQR2  
 100 36 59.0 82 16 Q93MT9  
 101 36 59.0 99 9 Q8SD72  
 102 36 59.0 124 6 Q7YS73  
 103 36 59.0 131 5 Q7YX99  
 104 36 59.0 217 15 Q8NTA9  
 105 36 59.0 227 15 Q90QX6  
 106 36 59.0 275 16 Q8F019  
 107 36 59.0 295 16 Q8F2V6  
 108 36 59.0 296 16 Q9ABP4  
 109 36 59.0 334 10 Q9FFP2  
 110 36 59.0 384 16 Q8F299  
 111 36 59.0 410 11 Q8BNT5  
 112 36 59.0 418 5 Q8YB1  
 113 36 59.0 418 11 Q8BNG8  
 114 36 59.0 439 16 Q8AIL2  
 115 36 59.0 442 16 Q9JXV8  
 116 36 59.0 442 16 Q9JW12  
 117 36 59.0 443 16 Q8ZS17  
 118 36 59.0 447 16 Q7V286  
 119 36 59.0 451 16 Q8DAD1  
 120 36 59.0 457 16 Q8Y206  
 121 36 59.0 462 16 Q9F022  
 122 36 59.0 462 16 Q8NTV4  
 123 36 59.0 463 16 Q8ZCA7  
 124 36 59.0 475 16 Q7WMN3  
 125 36 59.0 475 16 Q7WB66  
 126 36 59.0 500 5 Q9XUB5  
 127 36 59.0 512 10 Q9LUD3  
 128 36 59.0 856 16 Q8Y320  
 129 36 59.0 933 5 Q23360  
 130 36 59.0 1070 3 P78734  
 131 36 59.0 1115 10 Q9LSV7  
 132 36 59.0 1125 5 Q8I3M5  
 133 36 59.0 1275 16 Q99YA0  
 134 36 59.0 1275 16 Q8NZK6  
 135 36 59.0 1275 16 Q8K600  
 136 36 59.0 1277 16 Q879K2  
 137 35 57.4 49 5 Q7YVY9  
 138 35 57.4 60 5 Q7YVW9  
 139 35 57.4 72 15 Q72102  
 140 35 57.4 91 15 Q76445  
 141 35 57.4 99 15 Q9IH00  
 142 35 57.4 102 16 Q8FFV7  
 143 35 57.4 139 15 Q8AKJ7  
 144 35 57.4 148 2 Q9LS80  
 145 35 57.4 148 2 Q845X5  
 146 35 57.4 158 2 Q8GHP4  
 147 35 57.4 179 16 Q8G403  
 148 35 57.4 183 16 Q87633  
 149 35 57.4 201 8 Q9T227  
 150 35 57.4 220 2 Q87952

## ALIGNMENTS

RESULT 1  
 Q8BXH2 PRELIMINARY; PRT; 250 AA.  
 ID Q8BXH2  
 AC Q8BXH2;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similar to MVC box dependent interacting protein 1.  
 GN AMPH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 DR EMBL; AK047144; BAC32971.1; -.  
 DR MGD; MGI:103574; Amph.  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR SMART; SM00721; BAR; 1.  
 DR SMART; SM00721; BAR; 1.  
 SQ SEQUENCE 250 AA; 29125 MW; 568A3D55CC6C37CE CRC64;  
 Query Match 100.0%; Score 61; DB 11; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 0.013; 0; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RAQEKVLQKGA 13  
 DB 19 RAQEKVLQKGA 31  
 RESULT 2  
 Q8NPL3 PRELIMINARY; PRT; 385 AA.  
 ID Q8NPL3  
 AC Q8NPL3;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Amphiphysin I variant NT2 (Fragment).  
 GN AMPH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=22017878; PubMed=12023042;  
 RA Terada Y., Teutsui K., Sano K., Hosoya O., Ohtsuki H., Tokunaga A.,  
 RA Tsutsui K.;  
 RT "Novel splice variants of amphiphysin I are expressed in retina.";  
 RL FEBS Lett. 519:185-190 (2002).  
 DR EMBL; AF498097; AAM44811.1; -.  
 DR GO; GO:0006897; P: endocytosis; IEA.  
 DR GO; GO:0007268; P: synaptic transmission; IEA.  
 DR InterPro; IPR003005; Amphiphysin.  
 DR InterPro; IPR003017; Amphiphysin\_1.  
 DR InterPro; IPR006632; BAR.  
 DR InterPro; IPR004148; BAR\_dom.  
 DR Pfam; PF03114; BAR; 1.  
 DR PRINTS; PR01251; AMPHIPHYSIN.  
 DR ProDom; PD003208; Amphiphysin\_1; 1.  
 DR SMART; SM00721; BAR; 1.  
 FT NON\_TER 385 385  
 SQ SEQUENCE 385 AA; 43753 MW; 4B9AB0A3136711AA CRC64;  
 Query Match 100.0%; Score 61; DB 4; Length 385;  
 Best Local Similarity 100.0%; Pred. No. 0.02; 0; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 RAQEKVLQKLGKA 13
DB 19 RAQEKVLQKLGKA 31

RESULT 3
Q9BTH3 ID Q9BTH3 PRELIMINARY; PRT; 439 AA.
AC Q9BTH3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Similar to bridging integrator 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Kainine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator(TM) System Donor
vector."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; BC004101; AA04101.1; -.
DR EMBL; BT006865; AAP35511.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR InterPro; IPR000875; Cectoplin.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR PROSITE; PS00268; CECROPIN; 1.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 439 AA; 48286 MW; 350E428C8E198136 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RAQEKVLQKLGKA 13
DB 24 RAQEKVLQKLGKA 36

RESULT 4
Q7ZWP5 ID Q7ZWP5 PRELIMINARY; PRT; 478 AA.
AC Q7ZWP5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to bridging integrator 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Query Match 100.0%; Score 61; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OX Xenopodinae; Xenopus.
OX NCBI_TaxID=8395;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046852; AAH46852.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR PROSITE; PS00452; SH3DOMAIN.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
SQ SEQUENCE 478 AA; 53445 MW; 042A1A94D59095A8 CRC64;

Query Match 100.0%; Score 61; DB 13; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RAQEKVLQKLGKA 13
DB 23 RAQEKVLQKLGKA 35

RESULT 5
Q8WH9 ID Q8WH9 PRELIMINARY; PRT; 490 AA.
AC Q8WH9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amphiphysin IIB-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Kim K.-C., Kim T.-S., Kang K.-H., Choi K.-H.;
RT "Amphiphysin IIB-1, a novel splicing variant of amphiphysin II,
regulates p73b function through protein-protein interactions."
RL Oncogene 0:0-0(2002).
DR EMBL; AF411606; AAL38509.1; -.
DR GO; GO:0006897; P: endocytosis; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR003005; Amphiphysin.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01251; AMPHIPHYSIN.
DR PROSITE; PS00268; CECROPIN; 1.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
SQ SEQUENCE 490 AA; 53167 MW; 8F50F36F7B6E9690 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RAQEKVLQKLGKA 13
DB 24 RAQEKVLQKLGKA 36

```



LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-435-454-4

Query Match 100.0%; Score 61; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13  
DB 21 RAQEKVLQKLGKA 33

## RESULT 3

US-08-652-972A-4

Sequence 4, Application US/08652972A

Patent No. 5723581

GENERAL INFORMATION: ISSUED 3/2/1998

APPLICANT: Prendergast, George C.

APPLICANT: Sakamuro, Daitoku

TITLE OF INVENTION: Murine and Human Box-Dependent

TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P O Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/652,972A

FILING DATE: 24-MAY-1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,454

FILING DATE: 05-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: WST60BUSA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 451 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-652-972A-4

Query Match

Best Local Similarity 100.0%; Score 61; DB 1; Length 451;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAQEKVLQKLGKA 13

DB 21 RAQEKVLQKLGKA 33

## RESULT 4

US-08-919-145-6

Sequence 6, Application US/08919145

Patent No. 5958753

GENERAL INFORMATION:

APPLICANT: Prendergast, George C.  
TITLE OF INVENTION: Bau, A Binl Interacting Protein, and  
TITLE OF INVENTION: Uses Therefor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P.O. Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/919,145

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/025,482

FILING DATE: 23-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kodroff, Cathy A.

REGISTRATION NUMBER: 33,980

REFERENCE/DOCKET NUMBER: WST73AUSA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 451 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-919-145-6

Query Match

Best Local Similarity 100.0%; Score 61; DB 2; Length 451;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 RAQEKVLQKLGKA 13

DB

21 RAQEKVLQKLGKA 33

## RESULT 5

US-08-870-126-4

Sequence 4, Application US/08870126

Patent No. 6048702

GENERAL INFORMATION:

APPLICANT: Prendergast, George C.

APPLICANT: Sakamuro, Daitoku

TITLE OF INVENTION: Murine and Human Box-Dependent

TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P O Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/870,126

FILING DATE:

CLASSIFICATION: 530

101 35.5 37.4 724 4 US-09-562-737-23  
 102 35.5 37.4 916 4 US-08-252-991A-23637  
 103 35.5 37.4 1441 4 US-09-252-991A-28143  
 104 35.5 37.4 12 3 US-09-060-039-22  
 105 35.5 36.8 54 2 US-08-592-646A-15  
 106 35.5 36.8 54 4 US-09-165-422-15  
 107 35.5 36.8 83 4 US-09-134-000C-4372  
 108 35.5 36.8 96 4 US-09-621-976-4959  
 109 35.5 36.8 139 4 US-08-107-532A-4424  
 110 35.5 36.8 153 3 US-09-154-083-12  
 111 35.5 36.8 153 3 US-09-154-083-20  
 112 35.5 36.8 181 4 US-09-134-000C-6757  
 113 35.5 36.8 218 4 US-09-252-991A-26105  
 114 35.5 36.8 222 4 US-09-543-681A-6244  
 115 35.5 36.8 257 4 US-09-134-001C-5042  
 116 35.5 36.8 283 4 US-08-107-532A-6360  
 117 35.5 36.8 318 4 US-09-107-532A-6731  
 118 35.5 36.8 325 4 US-09-540-236-2444  
 119 35.5 36.8 329 4 US-09-107-532A-7038  
 120 35.5 36.8 329 4 US-09-543-681A-4790  
 121 35.5 36.8 361 1 US-08-537-434-1  
 122 35.5 36.8 398 4 US-10-083-889-2  
 123 35.5 36.8 398 4 US-10-083-889-4  
 124 35.5 36.8 408 4 US-10-083-889-22  
 125 35.5 36.8 415 4 US-09-134-000C-3595  
 126 35.5 36.8 419 4 US-09-252-991A-30457  
 127 35.5 36.8 429 4 US-09-134-001C-4960  
 128 35.5 36.8 458 4 US-09-252-991A-22614  
 129 35.5 36.8 463 4 US-09-252-991A-24757  
 130 35.5 36.8 464 4 US-08-297-468-2  
 131 35.5 36.8 495 4 US-09-479-645A-2  
 132 35.5 36.8 495 4 US-09-479-645A-4  
 133 35.5 36.8 608 4 US-09-489-039A-12785  
 134 35.5 36.8 611 4 US-09-107-532A-4988  
 135 35.5 36.8 640 4 US-09-177-165A-30  
 136 35.5 36.8 675 1 US-08-386-495-10  
 137 35.5 36.8 675 5 PCT-US96-02331-10  
 138 35.5 36.8 699 4 US-09-808-701A-33  
 139 35.5 36.8 715 4 US-09-808-701A-33  
 140 35.5 36.8 742 4 US-09-500-123-12  
 141 35.5 36.8 771 4 US-09-462-284-2  
 142 35.5 36.8 771 4 US-09-079-592-2  
 143 35.5 36.8 811 4 US-09-500-123-9  
 144 35.5 36.8 811 4 US-09-540-236-2412  
 145 35.5 36.8 817 4 US-08-252-991A-20757  
 146 35.5 36.8 871 4 US-09-500-123-7  
 147 35.5 36.8 895 4 US-09-489-039A-13127  
 148 35.5 36.8 915 4 US-09-252-991A-24992  
 149 35.5 36.8 944 4 US-09-134-000C-5578  
 150 35.5 36.8 962 4 US-09-694-777A-24

ALIGNMENTS

RESULT 1  
 US-08-630-915A-24  
 ; Sequence/24, Application US/08630915A  
 ; Patent No. 6309820  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SPARKS, Andrew B.  
 ; APPLICANT: HOFFMAN, No. 6309820h  
 ; APPLICANT: KAY, Brian K.  
 ; APPLICANT: FOWLES, Dana M.  
 ; APPLICANT: MCCONNELL, Stephen J.  
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
 ; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
 ; TITLE OF INVENTION: USING SAME  
 ; NUMBER OF SEQUENCES: 227  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York

STATE: New York  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA: US/08/630,915A  
 APPLICATION NUMBER: US/08/630,915A  
 FILING DATE: 03-APR-1996  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Misrock, S. Leslie  
 REGISTRATION NUMBER: 18,872  
 REFERENCE/DOCKET NUMBER: 1101-174  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-8864/9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 24:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 404 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 US-08-630-915A-24

Query Match 100.0%; Score 95; DB 4; Length 404;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIAGKRLVDYDSARH 18  
 Db 104 RIAGKRLVDYDSARH 121

RESULT 2  
 US-08-630-915A-22  
 ; Sequence/22, Application US/08630915A  
 ; Patent No. 6309820  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SPARKS, Andrew B.  
 ; APPLICANT: HOFFMAN, No. 6309820h  
 ; APPLICANT: KAY, Brian K.  
 ; APPLICANT: FOWLES, Dana M.  
 ; APPLICANT: MCCONNELL, Stephen J.  
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
 ; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
 ; TITLE OF INVENTION: USING SAME  
 ; NUMBER OF SEQUENCES: 227  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA: US/08/630,915A  
 ; APPLICATION NUMBER: US/08/630,915A  
 ; FILING DATE: 03-APR-1996  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Misrock, S. Leslie  
 ; REGISTRATION NUMBER: 18,872  
 ; REFERENCE/DOCKET NUMBER: 1101-174  
 ; TELECOMMUNICATION INFORMATION:

B

102 (e)

B

TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-630-915A-22

Query Match 100.0%; Score 95; DB 4; Length 434;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIARGRKLVYDSARHH 18  
Db 139 RIARGRKLVYDSARHH 156

RESULT 3  
US-08-435-454-4  
Sequence 4, Application US/08435454  
Patent No. 5605830  
GENERAL INFORMATION:  
APPLICANT: Prendergast, George C.  
APPLICANT: Sakamuro, Daitoku  
TITLE OF INVENTION: Murine and Human C-Myc Interacting  
TITLE OF INVENTION: Protein and Uses Therefor  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,454  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-435-454-4

Query Match 100.0%; Score 95; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIARGRKLVYDSARHH 18  
Db 136 RIARGRKLVYDSARHH 153

RESULT 4  
US-08-652-972A-4

Sequence 4, Application US/08652972A  
Patent No. 5723581  
GENERAL INFORMATION:  
APPLICANT: Prendergast, George C.  
APPLICANT: Sakamuro, Daitoku  
TITLE OF INVENTION: Murine and Human Box-Dependent  
TITLE OF INVENTION: MYC-Interacting Protein (BIN1) and Uses Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,972A  
FILING DATE: 24-MAY-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,454  
FILING DATE: 05-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-972A-4

Query Match 100.0%; Score 95; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIARGRKLVYDSARHH 18  
Db 136 RIARGRKLVYDSARHH 153

RESULT 5  
US-08-919-145-6  
Sequence 6, Application US/08919145  
Patent No. 5958753  
GENERAL INFORMATION:  
APPLICANT: Prendergast, George C.  
APPLICANT: Bau, A Binl Interacting Protein, and  
TITLE OF INVENTION: Uses Therefor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: